INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

An a posteriori error estimate of hierarchical type

R. Araya & P. Le Tallec

No 3722

June 24, 1999

THÈME 4

Rapport de recherche

ISSN 0249-6399

An *a posteriori* error estimate of hierarchical type

R. Araya* & P. Le Tallec†

Thème 4 — Simulation et optimisation
de systèmes complexes
Projet M3N

Rapport de recherche n ° 3722 — June 24, 1999 — 48 pages

Abstract: In this work we present a new a posteriori parameter free error estimate of hierarchical type and we apply this error estimate to the elasticity equations for linear isotropic heterogeneous materials. This estimate is proved to be optimal, independently of the material heterogeneities. The proof introduces an easy to check fundamental inf-sup condition, which implies the saturation assumption and partition lemma usually assumed in the analysis of hierarchical estimates. Insight on the numerical implementation and various numerical examples are finally presented .

Key-words: *A posteriori* error estimate, linear elasticity, bubble functions, heterogeneity.

(Résumé : tsvp)

* INRIA, Domaine de Voluceau- Rocquencourt- B.P. 105- Le Chesnay Cedex (France) `e-mail: Rodolfo.Araya@inria.fr`.

† Université Paris-Dauphine, CEREMADE, 75 7756 - Paris Cedex 16 (France) `e-mail: letallec@ceremade.dauphine.fr`

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
Téléphone : 01 39 63 55 11 - International : +33 1 39 63 55 11
Télécopie : (33) 01 39 63 53 30 - International : +33 1 39 63 53 30

Un estimateur d'erreur *a posteriori* de type hierarchique

Résumé : Dans ce travail on présente un nouvel estimateur d'erreur a posteriori de type hiérarchique pour les problèmes d'élasticité linéaire hétérogène isotrope. Cet estimateur s'applique dans un cadre très général et il est totalement intrinsèque. Nous démontrons son optimalité en nous fondant sur une condition inf-sup, assez facile à vérifier et permettant de vérifier l'hypothèse de saturation et le lemme de partition utilisés dans l'analyse classique. Cet article se conclut par une description de l'implémentation numérique et par quelques validations numériques.

Mots-clé : Estimateur d'erreur *a posteriori*, problèmes elliptiques fortement hétérogènes, fonctions bulles.

Contents

1 Introduction **4**

2 Model problem and notation **6**
2.1 The continuous problem 6
2.2 Finite element discretization 7
2.3 Edges and vertices 8
2.4 Bubble functions 9

3 Hierarchical intrinsic error estimator **11**
3.1 Introduction 11
3.2 Abstract Construction and fundamental example 11
3.3 A particular case 13

4 Analysis of the hierarchical estimate **16**
4.1 Hierarchical residual subspaces 16
4.2 Hierarchical approach. First step. 21
4.3 Verification of the saturation assumption and partition lemma . 25
4.4 Direct Analysis 29

5 Numerical implementation **31**
5.1 Computation of $a(P_i\mathbf{e}, P_i\mathbf{e})$ 31

6 Numerical examples **38**

7 Conclusion **47**

1 Introduction

When we use computers to simulate a physical phenomena, we have to face two sources of errors : modeling approximations and numerical errors. From the numerical analysis point of view, the relevant problem is the second one i.e. how to measure, to control and, hopefully, to minimize the error generated by the numerical approximation of the equations. The answer to these questions is based on the a posteriori error estimates introduced in the pioneering work of Babuska and Rheinboldt [3]. The idea is to give some information about the approximation error in a computable way which is something completely different from the so called a priori estimates. A posteriori error estimates, applied to the resolution of partial differential equations, are now very well known and established, see [1] for a survey on this topic.

In this paper we are interested in a generalization of the work made by Bank and Weiser (see [4]), on *hierarchical a posteriori error estimates*. In the framework of compressible heterogeneous linear elasticity, we propose a parameter free hierarchical estimate, well suited for industrial applications using complex geometries.

This paper is organized as follows. In section 2 we introduce both continuous and discrete formulations of our model problem along with the definitions and properties of the standard bubble functions. In section 3 we give the definition of our new error estimate based on a hierarchical decomposition of the discrete subspace into generalized bubble functions subspaces, and we also provide an example of these generalized bubble functions. Finally we present a relation between our hierarchical error estimate and the residual based one proposed by Araya and Le Tallec in [2].

In section 4 we carry out the theoretical analysis of our error estimate in two steps proving the equivalence between the error estimate and the approximation error independently of the stiffness of the composing materials, but with constants possibly depending on the Poisson ratios and on the anisotropy of the materials. The first step, the standard one, is to prove this equivalence assuming a saturation hypothesis and a partition lemma. These conditions are, in general, not easy to prove. The second step uses, as starting point, an inf-sup condition, which in our particular case implies the saturation assumption and has the advantage of being easily fulfilled.

Finally sections 5 to 7 are devoted to the numerical implementation of the error estimate, to the description of various practical examples and conclusions.

2 Model problem and notation

2.1 The continuous problem

Let Ω a Lipschitz, bounded domain of $\mathbb{R}^N$ with an interior boundary denoted by Γ_{12}. This boundary represents the interface between two elastic, isotropic and homogeneous materials, noted Ω_1 and Ω_2, respectively (see Fig 1). Let $\Gamma := \partial\Omega$ be divided in two parts $\Gamma = \Gamma_D \cup \Gamma_N$, with $\Gamma_D \cap \Gamma_N = \emptyset$, $meas(\Gamma_D) \neq 0$ and $\partial\Omega_i \cap \Gamma_D \neq \emptyset$, $i = 1, 2$. The body is fixed on Γ_D, which means for the time being that we assume that each subdomain is fixed on part of its boundary. In this framework, we consider the following elasticity problem

$$(\mathrm{P}) \qquad \left\{ \begin{aligned} -\mathrm{div}\, \boldsymbol{\sigma} &= \mathbf{f} && \text{in } \Omega\,, \\ \mathbf{u} &= \mathbf{0} && \text{on } \Gamma_D\,, \\ \boldsymbol{\sigma} \cdot \mathbf{n} &= \mathbf{g} && \text{on } \Gamma_N\,, \end{aligned} \right.$$

where $\mathbf{f} \in \mathrm{L}^2(\Omega)^N$ and $\mathbf{g} \in \mathrm{L}^2(\Gamma_D)^N$ are the given external forces and $\boldsymbol{\sigma}$ is the stress tensor. This tensor satisfies the constitutive law

$$\boldsymbol{\sigma} = (\sigma_{ij}) = \left\{ \begin{array}{ll} \mathcal{A}_1 \boldsymbol{\varepsilon}(\mathbf{u}) & \text{on } \Omega_1\,, \\ \mathcal{A}_2 \boldsymbol{\varepsilon}(\mathbf{u}) & \text{on } \Omega_2\,, \end{array} \right.$$

with $\mathcal{A}_i$ the elasticity tensor of the constitutive material Ω_i and $\epsilon_{ij}(\mathbf{u}) := \frac{1}{2}(u_{i,j} + u_{j,i})$ the components of the linearized strain tensor $\boldsymbol{\varepsilon}(\mathbf{u})$ associated to $\mathbf{u}$. We assume that the constitutive materials are compressible and of bounded anisotropy, which means that we have

$$CE(x)|\boldsymbol{\varepsilon}(\mathbf{u})|^2 \leq \mathcal{A}(x)\boldsymbol{\varepsilon}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{u}) \leq \tilde{C}E(x)|\boldsymbol{\varepsilon}(\mathbf{u})|^2 \tag{2.1}$$

where the positive constants C and $\tilde{C}$ are independent of an average Young modulus $E(x)$ defined for each material.

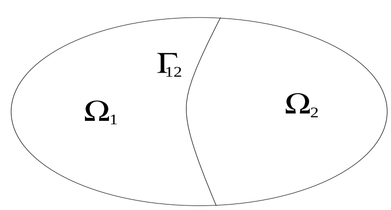

Figure 1: $\bar{\Omega} = \bar{\Omega}_1 \cup \bar{\Omega}_2$

The standard weak formulation of problem (P) is then:
Find $\mathbf{u} \in \mathbf{H}$ *such that*

$$a(\mathbf{u},\mathbf{v}) = < \mathbf{F}, \mathbf{v} > \quad , \forall \mathbf{v} \in \mathbf{H}, \tag{2.2}$$

where the space of admissible velocity fields and energy forms are given by

$$\begin{aligned} \mathbf{H} &:= \{\mathbf{v} \in \mathrm{H}^1(\Omega)^N \,/\, \mathbf{v} = \mathbf{0} \text{ on } \Gamma_D\}, \\ a(\mathbf{u},\mathbf{v}) &:= \int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{v}) = \int_\Omega \mathcal{A}\boldsymbol{\varepsilon}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{v}) \\ < \mathbf{F}, \mathbf{v} > &:= \int_\Omega \mathbf{f} \cdot \mathbf{v} + \int_{\Gamma_N} \mathbf{g} \cdot \mathbf{v} \,. \end{aligned}$$

We define the energy norm by:

$$\begin{aligned} \|\mathbf{v}\|_\Omega^2 &:= a(\mathbf{v},\mathbf{v}) \\ &= \int_\Omega \mathcal{A}\boldsymbol{\varepsilon}(\mathbf{v}) : \boldsymbol{\varepsilon}(\mathbf{v}) \qquad \forall \mathbf{v} \in \mathbf{H} \,. \end{aligned} \tag{2.3}$$

By Korn's inequality, since Γ_D has a non empty measure, there exists a constant C_Ω, depending only on the geometry of Ω, such that

$$\|\mathbf{v}\|_\Omega \le C_\Omega \|\boldsymbol{\varepsilon}(\mathbf{v})\|_{0,\Omega} \quad , \forall \mathbf{v} \in \mathbf{H}. \tag{2.4}$$

Above $\|\cdot\|_{m,\Omega}$ denotes the H^m norm on Ω.

2.2 Finite element discretization

Let h be a positive discretization parameter, and consider a triangulation $\mathcal{T}_h$ of $\bar{\Omega}$, that is a partition of $\bar{\Omega}$ into non degenerate triangles T (respectively tetrahedra in dimension 3), with diameters bounded by h, and such that each pair of elements T_1 and T_2 of $\mathcal{T}_h$ are either disjoint or share a vertex, an edge or a complete face. We denote by h_T the diameter of T, by ρ_T the diameter of the circle (respectively, sphere) inscribed in T and we set

$$\sigma_T := \frac{h_T}{\rho_T} \,.$$

We assume that the family of triangulations $(\mathcal{T}_h)_h$ is *shape regular*, i.e. there exists a constant σ, independent of h, such that

$$\sigma_T \leq \sigma \qquad , \forall T \in \mathcal{T}_h \,. \tag{2.5}$$

Let $\mathbf{H}_h$ the finite element space defined by

$$\mathbf{H}_h := \{\mathbf{v}_h \in \mathcal{C}(\Omega)^N \,/\, \mathbf{v}_h = \mathbf{0} \text{ on } \Gamma_D \,,\, \mathbf{v}_h|_T \in \mathbb{P}_k(T)^N \,, \forall T \in \mathcal{T}_h\} \quad (k \geq 1) \,,$$

with $\mathbb{P}_k$ the set of polynomials of degree at most k.

Then the approximate problem of (2.2) reads: *Find* $\mathbf{u}_h \in \mathbf{H}_h$ *such that*

$$a(\mathbf{u}_h, \mathbf{v}_h) = < \mathbf{F}, \mathbf{v}_h > \quad , \forall \mathbf{v}_h \in \mathbf{H}_h \,. \tag{2.6}$$

In what follows we use the following notation

$$\begin{aligned} a &\preceq b \Longleftrightarrow a \leq Cb \\ a &\simeq b \Longleftrightarrow a \preceq b \text{ and } b \preceq a \,, \end{aligned}$$

where the constant C is independent of h and all equivalent Young moduli E_i.

2.3 Edges and vertices

For any element $T \in \mathcal{T}_h$ we denote by $\mathcal{E}(T)$ and $\mathcal{N}(T)$ the set of its edges (faces in dimension 3) and vertices, respectively, and set (see [11])

$$\mathcal{E}_{h,\Omega} := \bigcup_{T \in \mathcal{T}_h} \mathcal{E}(T) \,.$$

We split the set of faces $\mathcal{E}_{h,\Omega}$ into internal, interfaces and boundary faces,

$$\mathcal{E}_{h,\Omega} = \mathcal{E}_h \cup \mathcal{E}_{12} \cup \mathcal{E}_N \cup \mathcal{E}_D \,,$$

with

$$\begin{aligned} \mathcal{E}_N &:= \{F \in \mathcal{E}_{h,\Omega} \,/\, F \subset \Gamma_N\} \quad , \quad \mathcal{E}_D := \{F \in \mathcal{E}_{h,\Omega} \,/\, F \subset \Gamma_D\} \,, \\ \mathcal{E}_{12} &:= \{F \in \mathcal{E}_{h,\Omega} \,/\, F \subset \Gamma_{12}\} \quad , \quad \mathcal{E}_h := \{F \in \mathcal{E}_{h,\Omega} \,/\, F \notin \mathcal{E}_D \cup \mathcal{E}_N \cup \mathcal{E}_{12}\} \,. \end{aligned}$$

Given a $F \in \mathcal{E}_{h,\Omega}$ we denote by $\mathcal{N}(F)$ the set of its vertices. For $T \in \mathcal{T}_h$ and $F \in \mathcal{E}_{h,\Omega}$ we define their neighborhoods

$$
\begin{aligned}
w_T &:= \bigcup_{\mathcal{E}(T)\cap\mathcal{E}(T')\neq\emptyset} T' \,, \; w_F := \bigcup_{F\in\mathcal{E}(T')} T' \,, \\
\tilde{w}_T &:= \bigcup_{\mathcal{N}(T)\cap\mathcal{N}(T')\neq\emptyset} T' \,, \; \tilde{w}_F := \bigcup_{\mathcal{N}(F)\cap\mathcal{N}(T')\neq\emptyset} T' \,.
\end{aligned}
$$

Remark Condition (2.5) implies that h_T/h_F, $T \in \mathcal{T}_h$, $F \in \mathcal{E}(T)$, and $h_T/h_{T'}$, $T, T' \in \mathcal{T}_h$, $\mathcal{N}(T) \cap \mathcal{N}(T') \neq \emptyset$, are bounded from below and from above by constants which only depend on σ, with h_F denoting the diameter of face F.

2.4 Bubble functions

For each element $T \in \mathcal{T}_h$ we can define the *element bubble* function b_T^* by

$$
b_T^* := (N+1)^{N+1} \prod_{i=1}^{N+1} \lambda_{T,i} \,. \tag{2.7}
$$

Above $\lambda_{T,j}(x)$ denote the j *barycentric coordinates* of the point x in T. Similarly, for each edge (face) $F \in \mathcal{E}_{h,\Omega}$, we can define the *edge (face) bubble* function

$$
b_F^* := N^N \prod_{i=1}^{N} \lambda_{T,i} \,, \tag{2.8}
$$

with $T \in \omega_F$.

The above definition of b_F^* assumes that, say for $N = 2$, in each triangle of ω_F, the edge F is associated to the vertices with local numbers 1 and 2.

By construction, we have the following properties of the bubble functions b_T^* and b_F^*.

Lemma 2.1 *Let $T \in \mathcal{T}_h$ and $F \in \mathcal{E}_{h,\Omega}$ be arbitrary, then*

$$
\operatorname{supp} b_T^* \subset T, \quad 0 \le b_T^* \le 1, \quad \max_{x\in T} b_T^*(x) = 1, \tag{2.9}
$$

$$
\operatorname{supp} b_F^* \subset w_F, \quad 0 \le b_F^* \le 1, \quad \max_{x\in F} b_F^*(x) = 1 \,. \tag{2.10}
$$

Moreover, using standard discrete norm equivalence arguments, we can prove (see [2])

Lemma 2.2 *The following estimate holds for any local function $f \in \mathbb{P}_{k-2}(T)$ and $g \in \mathbb{P}_{k-1}(F)$*

$$
\begin{aligned}
\int_T b_T^* f^2 &\simeq \|f\|_{0,2,T}^2 \,, \\
\int_F b_F^* g^2 &\simeq \|g\|_{0,2,F}^2 \,, \\
\int_T |\nabla(b_T^* f)|^2 &\preceq h_T^{-2} \int_T f^2 \,,
\end{aligned}
$$

$$
\begin{aligned}
\int_{T \in w_F} |\nabla(b_F^* g)|^2 &\preceq h_F^{-1} \int_F g^2 \,, \\
\int_{T \in w_F} |b_F^* g|^2 &\preceq h_F \int_F g^2 \,.
\end{aligned}
$$

We finally have the following technical result, which is a simple consequence of the basic finite element theory (see [7]).

Lemma 2.3 *Let $F \in \mathcal{E}_h \cup \mathcal{E}_{12}$ and $w_F = T_1 \cup T_2$, then*

$$
\|b_F^*\|_{0,F} \preceq h_F \|\nabla b_F^*\|_{0,T_i} \tag{2.11}
$$

$$
\|b_F^*\|_{0,T_i} \preceq h_{T_i} \|\nabla b_F^*\|_{0,T_i} \,. \tag{2.12}
$$

3 Hierarchical intrinsic error estimator

3.1 Introduction

We wish to introduce a parameter free optimal *a posteriori* error estimator. For this purpose, we apply the general theory developed by Bank and Weiser (cf. [4]), but with possible different choices of local spaces.

3.2 Abstract Construction and fundamental example

Consider a finer finite element space $\mathbf{W}_h$ such that $\mathbf{H}_h \subset \mathbf{W}_h \subset \mathbf{H}$ and denote by $\mathbf{w}_h$ the unique solution of problem (2.6) in $\mathbf{W}_h$. We assume that there exist M subspaces $\mathbf{H}_i$ of $\mathbf{W}_h$ such that

$$\mathbf{W}_h = \mathbf{H}_0 + \sum_{i=1}^{M} \mathbf{H}_i \,, \tag{3.1}$$

with $\mathbf{H}_0 = \mathbf{H}_h$.

Associated with each subspace $\mathbf{H}_i$ there is a projection operator $P_i : \mathbf{H} \to \mathbf{H}_i$ given by the solution of the local elasticity problem

$$a(P_i\mathbf{v}, \mathbf{w}_i) = a(\mathbf{v}, \mathbf{w}_i), \quad \forall \mathbf{w}_i \in \mathbf{H}_i \,, P_i\mathbf{v} \in \mathbf{H}_i \,.$$

Let finally $P_W : \mathbf{H} \to \mathbf{W}_h$ be the projection operator defined by the global elasticity problem

$$a(P_W\mathbf{v}, \mathbf{z}_h) = a(\mathbf{v}, \mathbf{z}_h) \quad \forall \mathbf{z}_h \in \mathbf{W}_h \,, P_W\mathbf{v} \in \mathbf{W}_h.$$

We obviously have $\mathbf{w}_h = P_W\mathbf{u}$.

In this framework, the error $\mathbf{e} = \mathbf{u} - \mathbf{u}_h$ in the approximation of the solution $\mathbf{u}$ of (2.2) by elements $\mathbf{u}_h$ of $\mathbf{H}_h$ satisfy

$$a(\mathbf{e}, \mathbf{v}) = < R_h, \mathbf{v} > \qquad , \forall \mathbf{v} \in \mathbf{H} \,, \tag{3.2}$$

where the *residual* R_h is the element of $\mathbf{H}'$ given by

$$< R_h, \mathbf{v} > := < \mathbf{F}, \mathbf{v} > - a(\mathbf{u}_h, \mathbf{v}) \qquad , \forall \mathbf{v} \in \mathbf{H} \,. \tag{3.3}$$

Using the Green's theorem we can represent the residual R_h by

$$< R_h, \mathbf{v} >= \sum_{T \in \mathcal{T}_h} (R_T, \mathbf{v})_{0,T} + \sum_{F \in \mathcal{E}_{h,\Omega}} (R_F, \mathbf{v})_{0,F} \tag{3.4}$$

where

$$R_T = (\mathrm{div}\, \boldsymbol{\sigma}_h + \mathbf{f})|_T \qquad , T \in \mathcal{T}_h \tag{3.5}$$

and

$$R_F = \begin{cases} \mathbf{0} & \text{if } F \in \mathcal{E}_D, \\ \mathbf{g} - \boldsymbol{\sigma}_h \cdot \mathbf{n} & \text{if } F \in \mathcal{E}_N, \\ -\boldsymbol{\sigma}_{h,T} \cdot \mathbf{n}_T - \boldsymbol{\sigma}_{h,T'} \cdot \mathbf{n}_{T'} & \text{if } F \in \mathcal{E}_h \cup \mathcal{E}_{12}. \end{cases} \tag{3.6}$$

With the above definitions our *hierarchical a posteriori error estimate* η_H can be defined by the local additive decomposition

$$\eta_H = \{\sum_{i=1}^{M} a(P_i\mathbf{e}, P_i\mathbf{e})\}^{1/2} . \tag{3.7}$$

We recall that $P_i\mathbf{e}$ is the solution of the local elasticity subproblem

Find $\mathbf{e}_i \in \mathbf{H}_i$ *such that*

$$a(\mathbf{e}_i, \mathbf{v}_i) = a(\mathbf{e}, \mathbf{v}_i) =< R_h, \mathbf{v}_i > \qquad , \forall \mathbf{v}_i \in \mathbf{H}_i \, . \tag{3.8}$$

This subproblem only involves the mechanical residual and the local elasticity operators.

In section 4 we will give some theoretical information about η_H. For the time being, we restrict our attention to the choice of the finite dimensional subspaces $\mathbf{H}_i$. The idea is to choose these subspaces as local as possible and thus to compute $P_i\mathbf{e}$ in a cheap way. As a basic example we will use local face and element subspaces (therefore different from Bank and Weiser) given by

$$H_{B_T} = span\{b_T^* R_T\} \qquad , T \in \mathcal{T}_h \, , \tag{3.9}$$

$$H_{B_F} = span\{b_F^* R_F\} \qquad , F \in \mathcal{E}_h \cup \mathcal{E}_{12} \cup \mathcal{E}_N \, , \tag{3.10}$$

where b_T^* and b_F^* are the element and face bubble functions defined in (2.7) and (2.8), respectively. Thus η_H can be written as

$$\eta_H = \{\sum_{T\in\mathcal{T}_h} a(P_T\mathbf{e}, P_T\mathbf{e}) + \sum_{F\in\mathcal{E}_h\cup\mathcal{E}_N\cup\mathcal{E}_{12}} a(P_F\mathbf{e}, P_F\mathbf{e})\}^{1/2}. \qquad (3.11)$$

General theoretical results on η_H will be proved later, both in a general framework and in the particular case given by (3.11).

3.3 A particular case

In the particular case when $\mathbf{H}_i$ is one dimensional (cf. [5] section 3) we have the following result

Lemma 3.1 *If there exists* $\mathbf{v}_i$ *such that* $\mathbf{H}_i = span\{\mathbf{v}_i\}$, $i = 1, \dots, M$, *then our error estimator* η_H *reduces to*

$$\eta_H = \{\sum_{i=1}^{M} \frac{< R_h, \mathbf{v}_i >^2}{a(\mathbf{v}_i, \mathbf{v}_i)}\}^{1/2}.$$

Proof. Since $\mathbf{H}_i = span\{\mathbf{v}_i\}$, $i = 1, \dots, M$, there exist coefficients γ_i such that $P_i\mathbf{e} = \gamma_i\mathbf{v}_i$, yielding

$$\gamma_i a(\mathbf{v}_i, \mathbf{v}_i) = a(P_i\mathbf{e}, \mathbf{v}_i) = a(\mathbf{e}, \mathbf{v}_i) =< R_h, \mathbf{v}_i > .$$

By identification, we obtain

$$\gamma_i = \frac{< R_h, \mathbf{v}_i >}{a(\mathbf{v}_i, \mathbf{v}_i)},$$

and hence

$$a(P_i\mathbf{e}, P_i\mathbf{e}) = \gamma_i^2 a(\mathbf{v}_i, \mathbf{v}_i) = \frac{< R_h, \mathbf{v}_i >^2}{a(\mathbf{v}_i, \mathbf{v}_i)}.$$

From this and the definition of η_H the result follows. $\square$

We can then show the relationship between the weighted residual error estimate proposed by Araya and Le Tallec in [2] and the above hierarchical error estimate.

Lemma 3.2 *For isotropic heterogeneous materials, let $\eta_{R,T}$ be the error estimator given by*

$$\eta_{R,T} := \left\{ \frac{h_T^2}{E_T}\|\mathbf{f} + \operatorname{div}\boldsymbol{\sigma}_h\|_{0,T}^2 + \sum_{F\in\mathcal{E}(T)\cap\mathcal{E}_N} \frac{h_F}{E_T}\|\mathbf{g} - \boldsymbol{\sigma}_h\cdot\mathbf{n}\|_{0,F}^2 + \sum_{F\in\mathcal{E}(T)\cap(\mathcal{E}_h\cup\mathcal{E}_{12})} \frac{\alpha(T,F)^2}{E_T} h_F\|[\boldsymbol{\sigma}_h\cdot\mathbf{n}]\|_{0,F}^2 \right\}^{1/2}, \tag{3.12}$$

where E_T is the restriction of Young's modulus to the element T, and

$$\alpha(T_i, F) := \frac{E_{T_i}}{E_{T_1} + E_{T_2}}$$

with $F \in \partial T_1 \cap \partial T_2$. Then, for a proper choice of subspaces $\mathbf{H}_i$, we have

$$\eta_H \preceq \{\sum_{T\in\mathcal{T}_h} \eta_{R,T}^2\}^{1/2}. \tag{3.13}$$

This result will be proved later in a much more general framework, nevertheless we give below a constructive proof which helps to understand and justify the residual error estimator given by (3.12).

Proof of Lemma 3.2. We define the subspace $\mathbf{W}_h$ by $\mathbf{W}_h = \mathbf{H}_h + \sum_{i=1}^M \mathbf{H}_i$ where the one dimensional subspaces $\mathbf{H}_i$ are spanned by the functions $\mathbf{b}_F = \sum_{j=1}^N b_F^* \vec{\mathbf{e}}_j$, with $F \in \mathcal{E}_h \cup \mathcal{E}_{12} \cup \mathcal{E}_N$.

If $F \in \mathcal{E}_h \cup \mathcal{E}_{12}$, then

$$\begin{aligned} \frac{< R_h, \mathbf{b}_F >}{a(\mathbf{b}_F, \mathbf{b}_F)^{1/2}} &= \frac{< \mathbf{F}, \mathbf{b}_F > - a(\mathbf{u}_h, \mathbf{b}_F)}{a(\mathbf{b}_F, \mathbf{b}_F)^{1/2}} \\ &= \frac{\int_\Omega \mathbf{f}\cdot\mathbf{b}_F - \int_\Omega \boldsymbol{\sigma}(\mathbf{u}_h) : \boldsymbol{\varepsilon}(\mathbf{b}_F)}{a(\mathbf{b}_F, \mathbf{b}_F)^{1/2}} \\ &= \sum_i \{ \frac{\int_{T_i} (\mathbf{f} + \operatorname{div}\boldsymbol{\sigma}_h)\cdot\mathbf{b}_F - \int_{\partial T_i} \boldsymbol{\sigma}_h\cdot\mathbf{n}\cdot\mathbf{b}_F}{a(\mathbf{b}_F, \mathbf{b}_F)^{1/2}} \} \end{aligned}$$

$$
\begin{aligned}
&= \sum_i \{ \frac{\int_{T_i}(\mathbf{f} + \operatorname{div} \boldsymbol{\sigma}_h) \cdot \mathbf{b}_F}{a(\mathbf{b}_F, \mathbf{b}_F)^{1/2}} \} + \frac{\int_F [\boldsymbol{\sigma}_h \cdot \mathbf{n}] \cdot \mathbf{b}_F}{a(\mathbf{b}_F, \mathbf{b}_F)^{1/2}} \\
&\preceq \sum_i \{ \frac{\|\mathbf{f} + \operatorname{div} \boldsymbol{\sigma}_h\|_{0,T_i} \|\mathbf{b}_F\|_{0,T_i}}{\sqrt{E_i} \|\boldsymbol{\varepsilon}(\mathbf{b}_F)\|_{0,T_i}} \} + \frac{\|[\boldsymbol{\sigma}_h \cdot \mathbf{n}]\|_{0,F} \|\mathbf{b}_F\|_{0,F}}{(E_1 \|\boldsymbol{\varepsilon}(\mathbf{b}_F)\|_{0,T_1}^2 + E_2 \|\boldsymbol{\varepsilon}(\mathbf{b}_F)\|_{0,T_2}^2)^{1/2}} \\
&\preceq \sum_i \{ \frac{h_{T_i}}{\sqrt{E_i}} \|\mathbf{f} + \operatorname{div} \boldsymbol{\sigma}_h\|_{0,T_i} \} + \frac{\sqrt{h_F}}{\sqrt{E_1 + E_2}} \|[\boldsymbol{\sigma}_h \cdot \mathbf{n}]\|_{0,F} .
\end{aligned}
$$

On the other hand, if $F \in \mathcal{E}_N$ then

$$
\begin{aligned}
\frac{< R_h, \mathbf{b}_F >}{a(\mathbf{b}_F, \mathbf{b}_F)^{1/2}} &= \frac{< \mathbf{F}, \mathbf{b}_F > - a(\mathbf{u}_h, \mathbf{b}_F)}{a(\mathbf{b}_F, \mathbf{b}_F)^{1/2}} \\
&= \frac{\int_\Omega \mathbf{f} \cdot \mathbf{b}_F + \int_{\Gamma_N} \mathbf{g} \cdot \mathbf{b}_F - \int_\Omega \boldsymbol{\sigma}(\mathbf{u}_h) : \boldsymbol{\varepsilon}(\mathbf{b}_F)}{a(\mathbf{b}_F, \mathbf{b}_F)^{1/2}} \\
&= \frac{\int_T (\mathbf{f} + \operatorname{div} \boldsymbol{\sigma}_h) \cdot \mathbf{b}_F + \int_F \mathbf{g} \cdot \mathbf{b}_F - \int_{\partial T} \boldsymbol{\sigma}_h \cdot \mathbf{n} \cdot \mathbf{b}_F}{a(\mathbf{b}_F, \mathbf{b}_F)^{1/2}} \\
&= \frac{\int_T (\mathbf{f} + \operatorname{div} \boldsymbol{\sigma}_h) \cdot \mathbf{b}_F}{a(\mathbf{b}_F, \mathbf{b}_F)^{1/2}} + \frac{\int_F (\mathbf{g} - \boldsymbol{\sigma}_h \cdot \mathbf{n}) \cdot \mathbf{b}_F}{a(\mathbf{b}_F, \mathbf{b}_F)^{1/2}} \\
&\preceq \frac{\|\mathbf{f} + \operatorname{div} \boldsymbol{\sigma}_h\|_{0,T} \|\mathbf{b}_F\|_{0,T}}{\sqrt{E_T} \|\boldsymbol{\varepsilon}(\mathbf{b}_F)\|_{0,T}} + \frac{\|\mathbf{g} - \boldsymbol{\sigma}_h \cdot \mathbf{n}\|_{0,F} \|\mathbf{b}_F\|_{0,F}}{\sqrt{E_T} \|\boldsymbol{\varepsilon}(\mathbf{b}_F)\|_{0,T}} \\
&\preceq \frac{h_T}{\sqrt{E_T}} \|\mathbf{f} + \operatorname{div} \boldsymbol{\sigma}_h\|_{0,T} + \frac{\sqrt{h_F}}{\sqrt{E_T}} \|\mathbf{g} - \boldsymbol{\sigma}_h \cdot \mathbf{n}\|_{0,F} .
\end{aligned}
$$

Then using Lemma 3.1, we obtain

$$
\begin{aligned}
\eta_H \preceq & \left\{ \sum_{T \in \mathcal{T}_h} \frac{h_T^2}{E_T} \|\mathbf{f} + \operatorname{div} \boldsymbol{\sigma}_h\|_{0,2,T}^2 + \sum_{F \in \mathcal{E}(T) \cap \mathcal{E}_N} \frac{h_F}{E_T} \|\mathbf{g} - \boldsymbol{\sigma}_h \cdot \mathbf{n}\|_{0,2,F}^2 \right. \\
& \left. + \sum_{F \in \mathcal{E}(T) \cap (\mathcal{E}_h \cup \mathcal{E}_{12})} \frac{\alpha(T,F)^2}{E_T} h_F \|[\boldsymbol{\sigma}_h \cdot \mathbf{n}]\|_{0,2,F}^2 \right\}^{1/2} . \qquad \square
\end{aligned}
$$

4 Analysis of the hierarchical estimate

4.1 Hierarchical residual subspaces

The first step in the analysis is to properly identify the basic mathematical properties which ensure the efficiency of our proposed approach. This requires first a few definitions.

Let $\mathcal{T}_h$ be a triangulation of Ω. We first introduce the finite dimensional subspaces $\mathbf{H}_R$ of $\mathbf{H}$, called the *residual subspaces*, defined by

$$\mathbf{H}_R = \begin{cases} \mathbf{H}_{R_T} & \text{for each } T \in \mathcal{T}_h, \\ \mathbf{H}_{R_F} & \text{for each } F \in \mathcal{E}_h \cup \mathcal{E}_{12} \cup \mathcal{E}_N, \end{cases} \tag{4.1}$$

under the notation

$$\begin{aligned} \mathbf{H}_{R_T} &= span\{R_T\} \qquad T \in \mathcal{T}_h \,, \\ \mathbf{H}_{R_F} &= span\{R_F\} \qquad F \in \mathcal{E}_h \cup \mathcal{E}_{12} \cup \mathcal{E}_N \,. \end{aligned}$$

We also choose finite dimensional subspaces $\mathbf{H}_B$ of $\mathbf{H}$ of *bubble functions* given by

$$\mathbf{H}_B = \begin{cases} \mathbf{H}_{B_T} & \text{for each } T \in \mathcal{T}_h, \\ \mathbf{H}_{B_F} & \text{for each } F \in \mathcal{E}_h \cup \mathcal{E}_{12} \cup \mathcal{E}_N, \end{cases} \tag{4.2}$$

where $\mathbf{H}_{B_T} \subset [H_0^1(T) \cap C^0(T)]^N$ and $\mathbf{H}_{B_F} \subset [H_0^1(\omega_F) \cap C^0(\omega_F)]^N$. We finally introduce the following constants

$$\theta_T = \frac{h_T}{\sqrt{E_T}} \qquad \text{if } T \in \mathcal{T}_h \,, \tag{4.3}$$

$$\theta_F = \begin{cases} \frac{\sqrt{h_F}}{\sqrt{E_T}} & \text{if } F \in \mathcal{E}_N \,, \\ \frac{\sqrt{h_F}}{\sqrt{E_{T_1}+E_{T_2}}} & \text{if } F \in \mathcal{E}_h \cup \mathcal{E}_{12} \,. \end{cases} \tag{4.4}$$

With these local subspaces we define the decomposition subspace $\mathbf{W}_h$ as in (3.1), i.e.

$$\mathbf{W}_h = \mathbf{H}_h + \sum_{T \in \mathcal{T}_h} \mathbf{H}_{B_T} + \sum_{F \in \mathcal{E}_h \cup \mathcal{E}_{12} \cup \mathcal{E}_N} \mathbf{H}_{B_F} = \mathbf{H}_0 + \sum_{i=1}^{M} \mathbf{H}_i \,. \tag{4.5}$$

The fundamental assumption on the local spaces $\mathbf{H}_R$ and $\mathbf{H}_B$ is the following *inf-sup condition*

(**LBB**) *There exist a positive constant β, independent of h and the physical coefficients, such that*

$$\inf_{\mathbf{r}\in\mathbf{H}_R} \sup_{\mathbf{b}\in\mathbf{H}_B} \frac{(\mathbf{b},\mathbf{r})_{0,\Omega_i}}{\theta_i\|\mathbf{r}\|_{0,\Omega_i}\|\mathbf{b}\|_{\tilde{\Omega}_i}} \geq \beta\,, \tag{4.6}$$

where Ω_i represents the element T of $\mathcal{T}_h$, if $\mathbf{H}_R = \mathbf{H}_{R_T}$, or the element F of $\mathcal{E}_h \cup \mathcal{E}_{12} \cup \mathcal{E}_N$ if $\mathbf{H}_R = \mathbf{H}_{R_F}$, and with $\tilde{\Omega}_i = T$ if $\mathbf{H}_R = \mathbf{H}_{R_T}$ and $\tilde{\Omega}_i = \omega_F$ if $\mathbf{H}_R = \mathbf{H}_{R_F}$.

The (LBB) condition combined with the definition of $P_T\mathbf{e}$ implies that

$$\begin{aligned}
\theta_T\|R_T\|_{0,T} &\leq \frac{1}{\beta}\sup_{\mathbf{b}_T\in\mathbf{H}_{B_T}} \frac{(\mathbf{b}_T,R_T)_{0,T}}{\|\mathbf{b}_T\|_T} \\
&= \frac{1}{\beta}\sup_{\mathbf{b}_T\in\mathbf{H}_{B_T}} \frac{< R_h,\mathbf{b}_T >}{\|\mathbf{b}_T\|_T} \\
&= \frac{1}{\beta}\sup_{\mathbf{b}_T\in\mathbf{H}_{B_T}} \frac{a(P_T\mathbf{e},\mathbf{b}_T)}{a(\mathbf{b}_T,\mathbf{b}_T)^{1/2}} \\
&\preceq a(P_T\mathbf{e},P_T\mathbf{e})^{1/2}\,.
\end{aligned} \tag{4.7}$$

On the other hand, if $F \in \mathcal{E}_h \cup \mathcal{E}_{12}$ we obtain

$$\begin{aligned}
\theta_F\|R_F\|_{0,F} &\leq \frac{1}{\beta}\sup_{\mathbf{b}_F\in\mathbf{H}_{B_F}} \frac{(\mathbf{b}_F,R_F)_{0,F}}{a(\mathbf{b}_F,\mathbf{b}_F)^{1/2}} \\
&= \frac{1}{\beta}\sup_{\mathbf{b}_F\in\mathbf{H}_{B_F}} \frac{< R_h,\mathbf{b}_F > - \sum_{T\in\omega_F}(R_T,\mathbf{b}_F)_{0,T}}{a(\mathbf{b}_F,\mathbf{b}_F)^{1/2}} \\
&\leq \frac{1}{\beta}\sup_{\mathbf{b}_F\in\mathbf{H}_{B_F}} \frac{a(P_F\mathbf{e},\mathbf{b}_F)}{a(\mathbf{b}_F,\mathbf{b}_F)^{1/2}} + \sup_{\mathbf{b}_F\in\mathbf{H}_{B_F}} \sum_{T\in\omega_F} \frac{\|R_T\|_{0,T}\|\mathbf{b}_F\|_{0,T}}{a(\mathbf{b}_F,\mathbf{b}_F)^{1/2}} \\
&\preceq a(P_F\mathbf{e},P_F\mathbf{e})^{1/2} + \sum_{T\in\omega_F}\theta_T\|R_T\|_{0,T}\,,
\end{aligned} \tag{4.8}$$

because for any $T \in \omega_F$ we have

$$\begin{aligned} \frac{\|\mathbf{b}_F\|^2_{0,T}}{a_T(\mathbf{b}_F,\mathbf{b}_F)} &= \frac{\int_T \mathbf{b}_F \cdot \mathbf{b}_F}{a_T(\mathbf{b}_F,\mathbf{b}_F)} \\ &\preceq \frac{\int_T \mathbf{b}_F \cdot \mathbf{b}_F}{\frac{E_T}{h_T^2}(\int_T \mathbf{b}_F \cdot \mathbf{b}_F)} \\ &\preceq \theta_T^2 . \end{aligned}$$

In the same way we can show that for all $F \in \mathcal{E}_N$

$$\theta_F \|R_F\|_{0,F} \preceq a(P_F\mathbf{e}, P_F\mathbf{e})^{1/2} + \sum_{T\in\omega_F} \theta_T \|R_T\|_{0,T} .$$

Therefore we have shown the following result

Lemma 4.1 *If (LBB) is true, then*

$$\begin{aligned} < R_h, \mathbf{v} > &\preceq \sum_{T\in\mathcal{T}_h} \|R_T\|_{0,T}\|\mathbf{v}\|_{0,T} + \sum_{F\in\mathcal{E}_h\cup\mathcal{E}_{12}\cup\mathcal{E}_N} \|R_F\|_{0,F}\|\mathbf{v}\|_{0,F} \\ &\preceq \sum_{T\in\mathcal{T}_h} a(P_T\mathbf{e}, P_T\mathbf{e})^{1/2}\theta_T^{-1}\|\mathbf{v}\|_{0,T} \\ &\quad + \sum_{F\in\mathcal{E}_h\cup\mathcal{E}_{12}\cup\mathcal{E}_N} [a(P_F\mathbf{e}, P_F\mathbf{e})^{1/2} + \sum_{T\in\omega_F} a(P_T\mathbf{e}, P_T\mathbf{e})^{1/2}]\theta_F^{-1}\|\mathbf{v}\|_{0,F} \end{aligned}$$

for all $\mathbf{v} \in \mathbf{H}$.

We can also show that the bubble subspaces defined in (3.9) and (3.10) satisfy the (LBB) condition. For this purpose, provided that $\mathbf{f}|_T$ belong to a given finite dimensional subspace for any $T \in \mathcal{T}_h$, we define the following norms on $span\{R_T\}$

$$\|\mathbf{r}\|_1 := \{\frac{1}{h_T^2}\int_T b_T^*\mathbf{r}\cdot\mathbf{r}\}^{1/2} \tag{4.9}$$

$$\|\mathbf{r}\|_2 := \{\frac{1}{h_T^2}\int_T \mathbf{r}\cdot\mathbf{r}\}^{1/2} \tag{4.10}$$

$$\|\mathbf{r}\|_3 := \{\frac{1}{E_T} a(b_T^*\mathbf{r}, b_T^*\mathbf{r})\}^{1/2} . \tag{4.11}$$

Using Lemma 2.1 we have clearly $\|\mathbf{r}|\|_1 \leq \|\mathbf{r}\|_2$. On the other hand, denoting by $R_{\hat{T}}$ the value of the residual once transported on the reference triangle, and since $span\{R_{\hat{T}}\}$ is finite dimensional, we have by equivalence of the norms on $span\{R_{\hat{T}}\}$

$$\begin{aligned}
\int_T \mathbf{r}\cdot\mathbf{r} &= \frac{meas(T)}{meas(\hat{T})}\int_{\hat{T}} \hat{\mathbf{r}}\cdot\hat{\mathbf{r}} \\
&\preceq \frac{meas(T)}{meas(\hat{T})}\int_{\hat{T}} b^*_{\hat{T}}\hat{\mathbf{r}}\cdot\hat{\mathbf{r}} \\
&\preceq \int_{\hat{T}} b^*_{\hat{T}}\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}\,.
\end{aligned}$$

The equivalence between $\|\cdot\|_1$ and $\|\cdot\|_3$ follows again from scaling arguments and the equivalence of norms in a finite dimensional subspace. More precisely, using the affine application $\mathcal{F}:\hat{T}\to T$, transforming the reference element $\hat{T}$ onto an arbitrary element T,

$$\mathcal{F}(\hat{x}) = B\hat{x} + b\,,$$

and (2.1), we obtain

$$\begin{aligned}
a(b^*_T\mathbf{r}, b^*_T\mathbf{r}) &\simeq E_T\int_T \boldsymbol{\varepsilon}(b^*_T\mathbf{r}) : \boldsymbol{\varepsilon}(b^*_T\mathbf{r})\,dx \\
&\simeq \frac{1}{4}E_T|J_{\mathcal{F}^{-1}}|\int_{\hat{T}} |\hat{\nabla}(b^*_{\hat{T}}\hat{\mathbf{r}})B^{-1} + B^{-t}\hat{\nabla}(b^*_{\hat{T}}\hat{\mathbf{r}})^t|^2\,d\hat{x} \\
&\simeq E_T|J_{\mathcal{F}^{-1}}|\,h_T^{-2}\int_{\hat{T}} |\hat{\boldsymbol{\varepsilon}}(b^*_{\hat{T}}\hat{\mathbf{r}})|^2\,d\hat{x}\,.
\end{aligned}$$

Since $span\{R_{\hat{T}}\}$ is finite dimensional by construction of $\mathbf{f}_T$ and since $\int_{\hat{T}} b^*_{\hat{T}}\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}\,d\hat{x}$ and $\int_{\hat{T}}|\hat{\boldsymbol{\varepsilon}}(b^*_{\hat{T}}\hat{\mathbf{r}})|^2\,d\hat{x}$ are two norms on $span\{R_{\hat{T}}\}$, we have

$$\int_{\hat{T}} |\hat{\boldsymbol{\varepsilon}}(b^*_{\hat{T}}\hat{\mathbf{r}})|^2\,d\hat{x} \simeq \int_{\hat{T}} b^*_{\hat{T}}\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}\,d\hat{x}\,, \tag{4.12}$$

and thus

$$a(b^*_T\mathbf{r}, b^*_T\mathbf{r}) \simeq E_T|J_{\mathcal{F}^{-1}}|\,h_T^{-2}\int_{\hat{T}} b^*_{\hat{T}}\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}\,d\hat{x}\,. \tag{4.13}$$

Transporting the last integral back to the triangle T, we finally deduce that

$$a(b_T^*\mathbf{r}, b_T^*\mathbf{r}) \simeq E_T h_T^{-2} \int_T b_T^*\mathbf{r}\cdot\mathbf{r}\,dx\,.$$

Thus, from the definition of $\mathbf{H}_{B_T}$ ($b_T^*\mathbf{r} \in \mathbf{H}_{B_T} = span\{b_T^*\mathbf{R}_T\}$), we get

$$\begin{aligned}
\inf_{\mathbf{r}\in\mathbf{H}_{R_T}} \sup_{\mathbf{b}_T\in\mathbf{H}_{B_T}} \frac{(\mathbf{r},\mathbf{b}_T)_{0,T}}{\theta_T\|\mathbf{r}\|_{0,T}\, a(\mathbf{b}_T,\mathbf{b}_T)^{1/2}} &\geq \inf_{\mathbf{r}\in\mathbf{H}_{R_T}} \frac{(\mathbf{r}, b_T^*\mathbf{r})_{0,T}}{\theta_T\|\mathbf{r}\|_{0,T}\, a(b_T^*\mathbf{r}, b_T^*\mathbf{r})^{1/2}}\\
&= \inf_{\mathbf{r}\in\mathbf{H}_{R_T}} \frac{\|\mathbf{r}\|_1^2}{\|\mathbf{r}\|_2\,\|\mathbf{r}\|_3}\\
&\geq \beta\,.
\end{aligned}$$

The same analysis can be carried out for a face $F \in \mathcal{E}_h \cup \mathcal{E}_{12} \cup \mathcal{E}_N$. In fact, for $F \in \mathcal{E}_h \cup \mathcal{E}_{12}$, we can define the following norms on $span\{R_F\}$

$$\|\mathbf{r}\|_1 := \{\frac{1}{h_F}\int_F b_F^*\mathbf{r}\cdot\mathbf{r}\}^{1/2} \tag{4.14}$$

$$\|\mathbf{r}\|_2 := \{\frac{1}{h_F}\int_F \mathbf{r}\cdot\mathbf{r}\}^{1/2} \tag{4.15}$$

$$\|\mathbf{r}\|_3 := \{\frac{1}{E_{T_1}+E_{T_2}} a(b_F^*\mathbf{r}, b_F^*\mathbf{r})\}^{1/2} \tag{4.16}$$

where $\omega_F = T_1 \cup T_2$. Thus we have, using once again Lemma 2.1 and transporting to the reference element, that $\|\cdot\|_1 \simeq \|\cdot\|_2$. On the other hand, using the same arguments as above, we have

$$\begin{aligned}
\|\mathbf{r}\|_3^2 &= \frac{1}{E_{T_1}+E_{T_2}} a(b_F^*\mathbf{r}, b_F^*\mathbf{r})\\
&\simeq \frac{E_{T_1}}{E_{T_1}+E_{T_2}}\int_{T_1}|\boldsymbol{\varepsilon}(b_F^*\mathbf{r})|^2\,dx + \frac{E_{T_2}}{E_{T_1}+E_{T_2}}\int_{T_2}|\boldsymbol{\varepsilon}(b_F^*\mathbf{r})|^2\,dx\\
&\leq \int_{T_1}|\boldsymbol{\varepsilon}(b_F^*\mathbf{r})|^2\,dx + \int_{T_2}|\boldsymbol{\varepsilon}(b_F^*\mathbf{r})|^2\,dx\\
&\preceq \frac{meas(T_1)}{meas(\hat{T})} h_{T_1}^{-2}\int_{\hat{T}}|b_{\hat{F}}^*\hat{\mathbf{r}}|^2\,d\hat{x} + \frac{meas(T_2)}{meas(\hat{T})} h_{T_2}^{-2}\int_{\hat{T}}|b_{\hat{F}}^*\hat{\mathbf{r}}|^2\,d\hat{x}
\end{aligned}$$

$$\begin{aligned}
&\preceq \frac{meas(T_1)}{meas(\hat{T})} h_{T_1}^{-2} \int_{\hat{F}} |b_{\hat{F}}^* \hat{\mathbf{r}}|^2 \, d\hat{s} + \frac{meas(T_2)}{meas(\hat{T})} h_{T_2}^{-2} \int_{\hat{F}} |b_{\hat{F}}^* \hat{\mathbf{r}}|^2 \, d\hat{s} \\
&\preceq h_{T_1}^{-2} \, h_F \int_F |\mathbf{r}|^2 + h_{T_2}^{-2} \, h_F \int_F |\mathbf{r}|^2 \\
&\preceq h_F^{-1} \int_F |\mathbf{r}|^2 \\
&\preceq \|\mathbf{r}\|_2^2 \, .
\end{aligned}$$

Thus using the equivalence between $\|\cdot\|_1$ and $\|\cdot\|_2$ we obtain that $\|\cdot\|_3 \preceq \|\cdot\|_1$ and then

$$\begin{aligned}
\inf_{\mathbf{r}\in\mathbf{H}_{R_F}} \sup_{\mathbf{b}_F\in\mathbf{H}_{B_F}} \frac{(\mathbf{r}, \mathbf{b}_T)_{0,F}}{\theta_F \|\mathbf{r}\|_{0,F} \, a(\mathbf{b}_F, \mathbf{b}_F)^{1/2}} &\geq \inf_{\mathbf{r}\in\mathbf{H}_{R_F}} \frac{(\mathbf{r}, b_F^*\mathbf{r})_{0,F}}{\theta_F \|\mathbf{r}\|_{0,F} \, a(b_F^*\mathbf{r}, b_F^*\mathbf{r})^{1/2}} \\
&= \inf_{\mathbf{r}\in\mathbf{H}_{R_F}} \frac{\|\mathbf{r}\|_1^2}{\|\mathbf{r}\|_2 \, \|\mathbf{r}\|_3} \\
&\geq \beta \, .
\end{aligned}$$

4.2 Hierarchical approach. First step.

The basic error analysis of a hierarchical approach is classically based on two major assumptions. The first one assumes that the approximate hierarchical solution $\mathbf{w}_h$ converges towards $\mathbf{u}$ more rapidly than $\mathbf{u}_h$. This can be expressed in terms of the following *saturation assumption*

(SA) *There exists a positive constant α, with $\alpha < 1$, such that*

$$\|\mathbf{u} - P_W \mathbf{u}\|_\Omega \leq \alpha \|\mathbf{u} - \mathbf{u}_h\|_\Omega \, .$$

Remark The drawback of this approach is the fact that assumption (SA) is not easy to prove, except in simple cases. See [10] for a different approach without the hypothesis (SA). But we will see later that this assumption is in fact a consequence of the fundamental LBB inequality.

The second basic assumption corresponds to a *partition lemma* and is the fundamental ingredient of any additive Schwarz technique. The corresponding assumption, also to be checked in the next section, is :

(PL) *For all* $\mathbf{z}_h \in \mathbf{W}_h$, *there exists a representation* $\mathbf{z}_h = \mathbf{v}_0 + \sum_{i=1}^{M} \mathbf{v}_i$ *with* $\mathbf{v}_0 \in \mathbf{H}_h$ *and* $\mathbf{v}_i \in \mathbf{H}_i$ *and a positive constant* C_0 *such that*

$$a(\mathbf{v}_0, \mathbf{v}_0) + \sum_{i=1}^{M} a(\mathbf{v}_i, \mathbf{v}_i) \le C_0\, a(\mathbf{z}_h, \mathbf{z}_h)\,.$$

With both assumptions, (SA) and (PL), we can first prove

Lemma 4.2

$$(\frac{1}{1+\alpha})\|P_W\mathbf{u} - \mathbf{u}_h\|_\Omega \le \|\mathbf{u} - \mathbf{u}_h\|_\Omega \le (\frac{1}{1-\alpha})\|P_W\mathbf{u} - \mathbf{u}_h\|_\Omega\,.$$

Proof. By a simple application of (SA) we have that

$$\begin{aligned}
\|\mathbf{u} - \mathbf{u}_h\|_\Omega &\le \|\mathbf{u} - P_W\mathbf{u}\|_\Omega + \|P_W\mathbf{u} - \mathbf{u}_h\|_\Omega \\
&\le \alpha\|\mathbf{u} - \mathbf{u}_h\|_\Omega + \|P_W\mathbf{u} - \mathbf{u}_h\|_\Omega \\
&\le \frac{1}{1-\alpha}\|P_W\mathbf{u} - \mathbf{u}_h\|_\Omega\,.
\end{aligned}$$

Conversely

$$\begin{aligned}
\|P_W\mathbf{u} - \mathbf{u}_h\|_\Omega &\le \|\mathbf{u} - P_W\mathbf{u}\|_\Omega + \|\mathbf{u} - \mathbf{u}_h\|_\Omega \\
&\le \alpha\|\mathbf{u} - \mathbf{u}_h\|_\Omega + \|\mathbf{u} - \mathbf{u}_h\|_\Omega \\
&\le (1+\alpha)\|\mathbf{u} - \mathbf{u}_h\|_\Omega\,,
\end{aligned}$$

and the lemma follows. □

The optimality of our proposed hierarchical approach is now quite easy to prove.

Theorem 4.1 *Under the assumptions (SA) and (PL), if* $\mathbf{e} := \mathbf{u} - \mathbf{u}_h$ *is the error in the approximation of the solution of* (2.2) *by elements of* $\mathbf{H}_h$, *then*

$$\|\mathbf{e}\|_\Omega \simeq [\sum_{i=1}^{M} \|P_i\mathbf{e}\|_\Omega^2]^{1/2} \tag{4.17}$$

with constants independent of M, h *and material heterogeneities.*

Proof. By construction, from the definition of $P_i\mathbf{e}$, using Cauchy-Schwarz and the fact that $P_0\mathbf{e} = \mathbf{0}$, we have

$$\begin{aligned}
\left[\sum_{i=0}^{M} a(P_i\mathbf{e}, P_i\mathbf{e})\right]^2 &= \left[\sum_{i=1}^{M} a(\mathbf{e}, P_i\mathbf{e})\right]^2 \\
&= \left[a(\mathbf{e}, \sum_{i=1}^{M} P_i\mathbf{e})\right]^2 \\
&\leq a(\mathbf{e},\mathbf{e})\, a(\sum_{i=1}^{M} P_i\mathbf{e}, \sum_{i=1}^{M} P_i\mathbf{e})\,. \qquad (4.18)
\end{aligned}$$

But, using Cauchy-Schwarz, we have

$$\begin{aligned}
a(\sum_{i=1}^{M} P_i\mathbf{e}, \sum_{i=1}^{M} P_i\mathbf{e}) &= \sum_{i=1}^{M}\sum_{j\in I_i} a(P_i\mathbf{e}, P_j\mathbf{e}) \\
&\leq \sum_{i=1}^{M}\sum_{j\in I_i} \{\frac{1}{2}a(P_i\mathbf{e}, P_i\mathbf{e}) + \frac{1}{2}a(P_j\mathbf{e}, P_j\mathbf{e})\} \\
&\leq C_{max}\sum_{i=1}^{M} a(P_i\mathbf{e}, P_i\mathbf{e}), \qquad (4.19)
\end{aligned}$$

where I_i denotes the set of spaces $\mathbf{H}_j$ which neighbor $\mathbf{H}_i$, i.e.

$$I_i = \{\, j \,/\, \exists\, \mathbf{v}_j \in \mathbf{H}_j \text{ and } \mathbf{v}_i \in \mathbf{H}_i \text{ such that } a(\mathbf{v}_i, \mathbf{v}_j) \neq 0\}\,,$$

and where C_{max} denotes the maximal number of neighbors,

$$C_{max} = \max\{card(I_l)\,/\,1 \leq l \leq M\}\,. \qquad (4.20)$$

Then, from (4.18) and (4.19), it's clear that

$$\sum_{i=1}^{M} a(P_i\mathbf{e}, P_i\mathbf{e}) \leq C_{max}\, a(\mathbf{e},\mathbf{e})\,. \qquad (4.21)$$

On the other hand, using assumption (PL), there exist elements $\mathbf{u}_i \in \mathbf{H}_i$ such that

$$\mathbf{u}_h - P_W\mathbf{u} = \sum_{i=0}^{M} \mathbf{u}_i \,,$$

with

$$\sum_{i=0}^{M} a(\mathbf{u}_i, \mathbf{u}_i) \leq C_0 a(\mathbf{u}_h - P_W\mathbf{u}, \mathbf{u}_h - P_W\mathbf{u}) \,.$$

Thus, by Cauchy-Schwarz, we obtain

$$\begin{aligned}
&a(\mathbf{u}_h - P_W\mathbf{u}, \mathbf{u}_h - P_W\mathbf{u}) = a(\mathbf{u}_h - P_W\mathbf{u}, \sum_{i=0}^{M} \mathbf{u}_i) \\
&= \sum_{i=0}^{M} a(P_i(\mathbf{u}_h - P_W\mathbf{u}), \mathbf{u}_i) \\
&\leq \left\{ \sum_{i=0}^{M} a(P_i(\mathbf{u}_h - P_W\mathbf{u}), P_i(\mathbf{u}_h - P_W\mathbf{u})) \right\}^{1/2} \cdot \left\{ \sum_{i=0}^{M} a(\mathbf{u}_i, \mathbf{u}_i) \right\}^{1/2} \\
&\leq C_0^{1/2} \left\{ \sum_{i=0}^{M} a(P_i(\mathbf{u}_h - P_W\mathbf{u}), P_i(\mathbf{u}_h - P_W\mathbf{u})) \right\}^{1/2} \{a(\mathbf{u}_h - P_W\mathbf{u}, \mathbf{u}_h - P_W\mathbf{u})\}^{1/2},
\end{aligned}$$

which yields

$$a(\mathbf{u}_h - P_W\mathbf{u}, \mathbf{u}_h - P_W\mathbf{u}) \leq C_0 \sum_{i=0}^{M} a(P_i(\mathbf{u}_h - P_W\mathbf{u}), P_i(\mathbf{u}_h - P_W\mathbf{u})) \,.$$

But $\mathbf{H}_i \subset \mathbf{W}_h$ and thus we have

$$\begin{aligned}
a(P_i(P_W\mathbf{u}), \mathbf{v}_i) &= a(P_W\mathbf{u}, \mathbf{v}_i) \\
&= a(\mathbf{u}, \mathbf{v}_i) \\
&= a(P_i\mathbf{u}, \mathbf{v}_i) \quad \forall \mathbf{v}_i \in \mathbf{H}_i \,.
\end{aligned}$$

This implies that $P_i \circ P_W = P_i$ and thus

$$P_i(\mathbf{u}_h - P_W\mathbf{u}) = P_i(\mathbf{u}_h - \mathbf{u}) = P_i\mathbf{e} \,.$$

On the other hand, by Lemma 4.2 we have

$$a(\mathbf{u}_h - P_W\mathbf{u}, \mathbf{u}_h - P_W\mathbf{u}) \geq (1-\alpha)^2\, a(\mathbf{e},\mathbf{e})\,.$$

Hence we finally get

$$(1-\alpha)^2 a(\mathbf{e},\mathbf{e}) \leq C_0 \sum_{i=1}^{M} a(P_i\mathbf{e}, P_i\mathbf{e})\,. \qquad \square$$

4.3 Verification of the saturation assumption and partition lemma

The difficult part in the previous analysis is to verify the saturation assumption and the partition lemma. In our particular case, the (LBB) condition implies the saturation assumption, as we will show in the Lemma 4.4. For this purpose, we first prove

Lemma 4.3 *If (LBB) is true then*

$$a(\mathbf{e},\mathbf{e}) \preceq \sum_{T\in\mathcal{T}_h} a(P_T\mathbf{e}, P_T\mathbf{e}) + \sum_{F\in\mathcal{E}_h\cup\mathcal{E}_{12}\cup\mathcal{E}_N} a(P_F\mathbf{e}, P_F\mathbf{e})\,. \tag{4.22}$$

Proof. From Lemma 4.1, we have that

$$\begin{aligned} < R_h, \mathbf{v} > \;&\preceq \sum_{T\in\mathcal{T}_h} a(P_T\mathbf{e}, P_T\mathbf{e})^{1/2}\theta_T^{-1}\|\mathbf{v}\|_{0,T} \\ &+ \sum_{F\in\mathcal{E}_h\cup\mathcal{E}_{12}\cup\mathcal{E}_N} [a(P_F\mathbf{e}, P_F\mathbf{e})^{1/2} + \sum_{T\in\omega_F} a(P_T\mathbf{e}, P_T\mathbf{e})^{1/2}]\theta_F^{-1}\|\mathbf{v}\|_{0,F}\,. \end{aligned}$$

Let now $\Pi_h : \mathbf{H} \to \mathbf{H}_h$ be a projection operator of Clement type (cf. [8]). Then applying the above inequality with $\mathbf{v} = \mathbf{e} - \Pi_h\mathbf{e}$, we deduce from Cauchy-Schwarz

$$\begin{aligned} a(\mathbf{e},\mathbf{e}) &= < R_h, \mathbf{e} > \\ &= < R_h, \mathbf{e} - \Pi_h\mathbf{e} > \\ &\preceq \sum_{T\in\mathcal{T}_h} a(P_T\mathbf{e}, P_T\mathbf{e})^{1/2}\theta_T^{-1}\|\mathbf{e} - \Pi_h\mathbf{e}\|_{0,T} \\ &\quad + \sum_{F\in\mathcal{E}_h\cup\mathcal{E}_{12}\cup\mathcal{E}_N} [a(P_F\mathbf{e}, P_F\mathbf{e})^{1/2} + \sum_{T\in\omega_F} a(P_T\mathbf{e}, P_T\mathbf{e})^{1/2}]\theta_F^{-1}\|\mathbf{e} - \Pi_h\mathbf{e}\|_{0,F} \end{aligned}$$

$$
\begin{aligned}
\preceq \quad & \{\sum_{T\in\mathcal{T}_h} a(P_T\mathbf{e},P_T\mathbf{e}) + \sum_{F\in\mathcal{E}_h\cup\mathcal{E}_{12}\cup\mathcal{E}_N} a(P_F\mathbf{e},P_F\mathbf{e})\}^{1/2}\cdot \\
& \{\sum_{T\in\mathcal{T}_h} \theta_T^{-2}\|\mathbf{e}-\Pi_h\mathbf{e}\|_{0,T}^2 + \sum_{F\in\mathcal{E}_h\cup\mathcal{E}_{12}\cup\mathcal{E}_N} \theta_F^{-2}\|\mathbf{e}-\Pi_h\mathbf{e}\|_{0,F}^2\}^{1/2}.
\end{aligned}
$$

But, as proved in [8], we have under assumption (2.1)

$$
\begin{aligned}
\sum_{T\in\mathcal{T}_h} \theta_T^{-2}\|\mathbf{e}-\Pi_h\mathbf{e}\|_{0,T}^2 &\preceq a(\mathbf{e},\mathbf{e}) \\
\sum_{F\in\mathcal{E}_h\cup\mathcal{E}_{12}\cup\mathcal{E}_N} \theta_F^{-2}\|\mathbf{e}-\Pi_h\mathbf{e}\|_{0,F}^2 &\preceq a(\mathbf{e},\mathbf{e}),
\end{aligned}
$$

with constants independent of the local values of the Young modulus. Therefore our lemma follows directly. $\square$

We can now prove

Lemma 4.4 *If (LBB) is true then the saturation lemma holds, meaning that, there exists a positive constant α such that $\alpha < 1$ and*

$$
\|\mathbf{u}-P_W\mathbf{u}\|_\Omega \le \alpha\|\mathbf{u}-\mathbf{u}_h\|_\Omega .
$$

Proof. Using the shape regularity of $\mathcal{T}_h$, we have that C_{max} given by (4.20) is bounded uniformly and then by construction, we have

$$
P_i\mathbf{e} = P_i(P_W\mathbf{u}-\mathbf{u}_h). \tag{4.23}
$$

Thus, we deduce that

$$
\begin{aligned}
\sum_{i=0}^M a(P_i\mathbf{e},P_i\mathbf{e}) &= \sum_{i=0}^M a(P_i(P_W\mathbf{u}-\mathbf{u}_h),P_i\mathbf{e}) \\
&= \sum_{i=0}^M a(P_W\mathbf{u}-\mathbf{u}_h,P_i\mathbf{e}) \\
&\le a(P_W\mathbf{u}-\mathbf{u}_h,P_W\mathbf{u}-\mathbf{u}_h)^{1/2} a(\sum_{i=0}^M P_i\mathbf{e},\sum_{i=0}^M P_i\mathbf{e})^{1/2}.
\end{aligned}
$$

Proceeding as in (4.19), we get

$$\sum_{i=0}^{M} a(P_i\mathbf{e}, P_i\mathbf{e}) \le a(P_W\mathbf{u}-\mathbf{u}_h, P_W\mathbf{u}-\mathbf{u}_h)^{1/2} C_{max}^{1/2} \sum_{i=0}^{M} a(P_i\mathbf{e}, P_i\mathbf{e})^{1/2},$$

that is

$$\{\sum_{i=0}^{M} a(P_i\mathbf{e}, P_i\mathbf{e})\}^{1/2} \le C_{max}^{1/2} a(P_W\mathbf{u}-\mathbf{u}_h, P_W\mathbf{u}-\mathbf{u}_h)^{1/2}. \tag{4.24}$$

On the other hand, $\mathbf{u} - P_W\mathbf{u}$ is a-orthogonal to $\mathbf{W}_h$, yielding

$$\begin{aligned} a(\mathbf{e},\mathbf{e}) &= a(\mathbf{u}-\mathbf{u}_h, \mathbf{u}-\mathbf{u}_h) \\ &= a(\mathbf{u}-P_W\mathbf{u}+P_W\mathbf{u}-\mathbf{u}_h, \mathbf{u}-P_W\mathbf{u}+P_W\mathbf{u}-\mathbf{u}_h) \\ &= a(\mathbf{u}-P_W\mathbf{u}, \mathbf{u}-P_W\mathbf{u}) + a(P_W\mathbf{u}-\mathbf{u}_h, P_W\mathbf{u}-\mathbf{u}_h). \end{aligned} \tag{4.25}$$

From equations (4.25) and (4.24) and Lemma 4.3 we obtain

$$\begin{aligned} a(\mathbf{u}-P_W\mathbf{u}, \mathbf{u}-P_W\mathbf{u}) &= a(\mathbf{e},\mathbf{e}) - a(P_W\mathbf{u}-\mathbf{u}_h, P_W\mathbf{u}-\mathbf{u}_h) \\ &\le a(\mathbf{e},\mathbf{e}) - \frac{1}{C_{max}} \sum_{i=0}^{M} a(P_i\mathbf{e}, P_i\mathbf{e}) \\ &\le (1 - \frac{1}{CC_{max}})\, a(\mathbf{e},\mathbf{e}). \qquad \square \end{aligned}$$

We are now also ready to prove that (PL) condition is also valid in our case.

Lemma 4.5 *For the space constructed in* (4.5)*, there exists a positive constant* C_0*, independent of* h *and of physical constants, such that for all* $\mathbf{v}_h \in \mathbf{W}_h$*, there exists* $\mathbf{v}_0 \in \mathbf{H}_h$ *and* $\mathbf{v}_i \in \mathbf{H}_i$ *with*

$$\mathbf{v} = \mathbf{v}_0 + \sum_{i=1}^{M} \mathbf{v}_i, \tag{4.26}$$

and

$$a(\mathbf{v}_0, \mathbf{v}_0) + \sum_{i=1}^{M} a(\mathbf{v}_i, \mathbf{v}_i) \le C_0 a(\mathbf{v},\mathbf{v}). \tag{4.27}$$

Proof. Let $\mathbf{v} \in \mathbf{W}_h$, then we can write

$$\mathbf{v} = \mathbf{v}_0 + \sum_{i=1}^{M} \mathbf{v}_i \,. \tag{4.28}$$

Let $T \in \mathcal{T}_h$. Using (4.28), we can write

$$\mathbf{v}|_T = \mathbf{v}_{0,T} - \mathbf{m}_{r,T} + \mathbf{m}_{r,T} + \sum_{i=1}^{M} \mathbf{v}_{i,T} \,, \tag{4.29}$$

where $\mathbf{m}_{r,T}$ is any rigid motion of element T. Let

$$\tilde{\mathbf{v}}_{0,T} := \mathbf{v}_{0,T} - \mathbf{m}_{r,T} \quad \text{and} \quad \tilde{\mathbf{H}}_{0,T} := \mathbf{H}_{0,T}/M_{r,T} \,,$$

with $M_{r,T}$ the subspace of rigid motions of T. We have then

$$\mathbf{v}|_T = \tilde{\mathbf{v}}_{0,T} + \mathbf{m}_{r,T} + \sum_{i=1}^{M} \mathbf{v}_{i,T} \,. \tag{4.30}$$

Due to the definition of subspaces $\mathbf{H}_i$, $i = 1, \dots, M$, we can prove that the set

$$I_T := \{ \, j \; / \; \mathbf{v}_{j,T} \neq \mathbf{0} \, \} \tag{4.31}$$

is finite and that its number of element $\{ \, card(I_T) \, / \; T \in \mathcal{T}_h \, \}$ is uniformly bounded. Also we have that the subspaces $\{\mathbf{H}_i\}_{i \geq 1}$ are in direct sum, i.e.

$$\mathbf{H}_i \cap \mathbf{H}_j = \{\mathbf{0}\} \qquad , i \neq j \,. \tag{4.32}$$

Thus, using the fact that $\mathbf{m}_{r,T}$ is a rigid body motion ($a(\mathbf{m}_{r,T}, \mathbf{v}) = 0 \,, \forall \mathbf{v}$) we have

$$\begin{aligned} a_T(\mathbf{v}, \mathbf{v}) &= a_T(\tilde{\mathbf{v}}_{0,T} + \sum_{i \in I_T} \mathbf{v}_{i,T}, \tilde{\mathbf{v}}_{0,T} + \sum_{i \in I_T} \mathbf{v}_{i,T}) \\ &\simeq E_T \int_T |\boldsymbol{\varepsilon}(\tilde{\mathbf{v}}_{0,T} + \sum_{i \in I_T} \mathbf{v}_{i,T})|^2 \, dx \\ &\simeq E_T \int_{\hat{T}} |\hat{\boldsymbol{\varepsilon}}(\tilde{\mathbf{v}}_{0,\hat{T}} + \sum_{i \in I_{\hat{T}}} \mathbf{v}_{i,\hat{T}})|^2 \, \|B^{-1}\|^2 \, |J_{\mathcal{F}^{-1}}| d\hat{x} \,. \end{aligned}$$

But $\int_{\hat{T}} |\hat{\boldsymbol{\varepsilon}}(\tilde{\mathbf{v}}_{0,\hat{T}} + \sum_{i \in I_{\hat{T}}} \mathbf{v}_{i,\hat{T}})|^2 \, d\hat{x}$ is a norm on $\tilde{\mathbf{H}}_{0,\hat{T}} \times \prod_{i \in I_{\hat{T}}} \mathbf{H}_{i,\hat{T}}$. Indeed, if

$$\boldsymbol{\varepsilon}(\tilde{\mathbf{v}}_{0,T} + \sum_{i \in I_T} \mathbf{v}_{i,T}) = 0$$

then

$$\tilde{\mathbf{v}}_{0,T} + \sum_{i \in I_T} \mathbf{v}_{i,T} = \mathbf{0}$$

since $\tilde{\mathbf{v}}_{0,T} + \sum_{i \in I_T} \mathbf{v}_{i,T}$ is not equal to a rigid motion by construction. From this and (4.32) we obtain

$$\tilde{\mathbf{v}}_{0,T} = \mathbf{0} \quad , \quad \mathbf{v}_{i,T} = \mathbf{0} \quad , \forall i \in I_T \, .$$

Thus, using the equivalence of norms on $\tilde{\mathbf{H}}_{0,\hat{T}} \times \prod_{i \in I_{\hat{T}}} \mathbf{H}_{i,\hat{T}}$ and adding the rigid motions of T, we obtain

$$\begin{aligned}
a_T(\mathbf{v}, \mathbf{v}) &\simeq E_T \int_{\hat{T}} |\hat{\boldsymbol{\varepsilon}}(\tilde{\mathbf{v}}_{0,\hat{T}})|^2 \|B^{-1}\|^2 \, |J_{\mathcal{F}^{-1}}| d\hat{x} \\
&\quad + \sum_{i \in I_{\hat{T}}} E_T \int_{\hat{T}} |\hat{\boldsymbol{\varepsilon}}(\mathbf{v}_{i,\hat{T}})|^2 \|B^{-1}\|^2 \, |J_{\mathcal{F}^{-1}}| d\hat{x} \\
&\simeq a_T(\tilde{\mathbf{v}}_{0,T}, \tilde{\mathbf{v}}_{0,T}) + \sum_{i \in I_T} a_T(\mathbf{v}_{i,T}, \mathbf{v}_{i,T}) \\
&= a_T(\mathbf{v}_{0,T}, \mathbf{v}_{0,T}) + \sum_{i \in I_T} a_T(\mathbf{v}_{i,T}, \mathbf{v}_{i,T}) \, , \quad \forall T \in \mathcal{T}_h \, .
\end{aligned}$$

The result follows by addition on all triangles. $\square$

4.4 Direct Analysis

The systematical hierarchical approach followed up to now can in fact be by-passed by a direct analysis without using the saturation assumption (SA) and the partition lemma (PL), as shown in the next theorem.

Theorem 4.2 *Suppose that (LBB) is true, then*

$$\|\mathbf{e}\|_\Omega \simeq \{ \sum_{T \in \mathcal{T}_h} a(P_T \mathbf{e}, P_T \mathbf{e}) + \sum_{F \in \mathcal{E}_h \cup \mathcal{E}_N} a(P_F \mathbf{e}, P_F \mathbf{e}) \}^{1/2} \, , \tag{4.33}$$

where the equivalence constants are independent of h and the physical coefficients.

Proof. The inequality

$$\sum_{T \in \mathcal{T}_h} a(P_T \mathbf{e}, P_T \mathbf{e}) + \sum_{F \in \mathcal{E}_h \cup \mathcal{E}_N} a(P_F \mathbf{e}, P_F \mathbf{e}) \preceq a(\mathbf{e}, \mathbf{e}) \tag{4.34}$$

can be proved as in (4.21). The reverse inequality is a direct consequence of Lemma 4.3. $\square$

5 Numerical implementation

5.1 Computation of $a(P_i\mathbf{e}, P_i\mathbf{e})$

Let $P_i : \mathbf{H} \to \mathbf{H}_i$ be a orthogonal projection where $dim\,\mathbf{H}_i = n_i$. On each subspace $\mathbf{H}_i$ we have to solve the following local elasticity subproblem:

$$a(P_i\mathbf{e}, \mathbf{v}_i) =< R_h, \mathbf{v}_i > \qquad , \forall \mathbf{v}_i \in \mathbf{H}_i\,. \tag{5.1}$$

Let $\mathcal{B} = \{\boldsymbol{\psi}_1, \dots, \boldsymbol{\psi}_{n_i}\}$ be a base of $\mathbf{H}_i$ (bubble functions). Then

$$\begin{aligned} a(P_i\mathbf{e}, \boldsymbol{\psi}_k) &= a(\mathbf{e}, \boldsymbol{\psi}_k) \qquad , 1 \le k \le n_i \\ &= < \mathbf{F}, \boldsymbol{\psi}_k > - a(\mathbf{u}_h, \boldsymbol{\psi}_k) \\ &= < R_h, \boldsymbol{\psi}_k > \;. \end{aligned}$$

But

$$P_i\mathbf{e} = \sum_{j=1}^{n_i} \alpha_j \boldsymbol{\psi}_j = \vec{\boldsymbol{\alpha}}^T \vec{\boldsymbol{\Psi}}$$

and thus we obtain the following very simple linear system

$$\mathbf{A}_i \vec{\boldsymbol{\alpha}} = \vec{\mathbf{R}}$$

where

$$\begin{aligned} \mathbf{A}_i &= (a_{jk}) \quad i = T \text{ or } F \qquad (5.2) \\ \vec{\boldsymbol{\alpha}} &= (\alpha_j) \qquad (5.3) \\ \vec{\mathbf{R}} &= (R_k) \qquad (5.4) \end{aligned}$$

with

$$a_{jk} = a(\boldsymbol{\psi}_j, \boldsymbol{\psi}_k) = \sum_{T \subset supp(\boldsymbol{\psi}_j) \cap supp(\boldsymbol{\psi}_k)} \int_T \sigma(\boldsymbol{\psi}_j) : \varepsilon(\boldsymbol{\psi}_k) \tag{5.5}$$

and

$$R_k =< R_h, \boldsymbol{\psi}_k >= \sum_{T \subset supp(\boldsymbol{\psi}_k)} \{ \int_T \mathbf{f} \cdot \boldsymbol{\psi}_k + \int_{\partial T \cap \Gamma_N} \mathbf{g} \cdot \boldsymbol{\psi}_k - \int_T \sigma(\mathbf{u}_h) : \varepsilon(\boldsymbol{\psi}_k) \}\,. \tag{5.6}$$

Once $\vec{\boldsymbol{\alpha}}$ is computed, we then have simply

$$a(P_i\mathbf{e}, P_i\mathbf{e}) = \vec{\boldsymbol{\alpha}}^T A_i \vec{\boldsymbol{\alpha}} \,. \tag{5.7}$$

To enter in more details, we introduce the basis scalar matrix

$$P = [p_1, p_2, \dots, p_{N_T}] \tag{5.8}$$

where $p_1, p_2, \dots, p_{N_T}$ are the N_T scalar shape functions. In 3D, we denote the corresponding vector matrix by

$$\mathbf{P} = \begin{pmatrix} P & 0 & 0 \\ 0 & P & 0 \\ 0 & 0 & P \end{pmatrix} . \tag{5.9}$$

We have then

$$[u_T] = \begin{pmatrix} P & 0 & 0 \\ 0 & P & 0 \\ 0 & 0 & P \end{pmatrix} \begin{pmatrix} u_{1T} \\ u_{2T} \\ u_{3T} \end{pmatrix} = \mathbf{P} \cdot \vec{\mathbf{u}}_T \tag{5.10}$$

where $\vec{\mathbf{u}}_T$ represents the value of $\mathbf{u}$ on the nodes of the element T.

We also introduce the gradient matrices

$$\begin{aligned} DP &= \begin{pmatrix} \frac{\partial p_1}{\partial x} & \frac{\partial p_2}{\partial x} & \dots & \frac{\partial p_n}{\partial x} \\ \frac{\partial p_1}{\partial y} & \frac{\partial p_2}{\partial y} & \dots & \frac{\partial p_n}{\partial y} \\ \frac{\partial p_1}{\partial z} & \frac{\partial p_2}{\partial z} & \dots & \frac{\partial p_n}{\partial z} \end{pmatrix} \\ \mathbf{DP} &= \begin{pmatrix} DP & 0 & 0 \\ 0 & DP & 0 \\ 0 & 0 & DP \end{pmatrix} . \end{aligned}$$

The gradient $\mathbf{Du}|_T$ of $\mathbf{u}$ on T, can now be expressed in function of $\mathbf{DP}$ and $\vec{\mathbf{u}}_T$ by

$$\mathbf{Du}|_T = \begin{pmatrix} \frac{\partial u_1}{\partial x} \\ \frac{\partial u_1}{\partial y} \\ \frac{\partial u_1}{\partial z} \\ \dots \\ \frac{\partial u_3}{\partial z} \end{pmatrix} = \mathbf{DP} \cdot \vec{\mathbf{u}}_T \,.$$

The tensors $\boldsymbol{\sigma}$ and $\boldsymbol{\varepsilon}$ can then be written as

$$\{\boldsymbol{\varepsilon}\} = \begin{pmatrix} \epsilon_{11} \\ \epsilon_{22} \\ \epsilon_{33} \\ 2\epsilon_{12} \\ 2\epsilon_{23} \\ 2\epsilon_{31} \end{pmatrix} \quad \text{and} \quad \{\boldsymbol{\sigma}\} = \begin{pmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \\ \sigma_{23} \\ \sigma_{31} \end{pmatrix}.$$

We can write $\{\boldsymbol{\varepsilon}\}$ like

$$\{\boldsymbol{\varepsilon}\} = \mathbf{D} \cdot \mathbf{DP} \cdot \vec{\mathbf{u}}_T$$

where

$$\mathbf{D} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

We can express $\{\boldsymbol{\sigma}\}$ by means of $\{\boldsymbol{\varepsilon}\}$

$$\{\boldsymbol{\sigma}\} = \mathcal{A} \cdot \{\boldsymbol{\varepsilon}\} = \mathcal{A} \cdot \mathbf{D} \cdot \mathbf{DP} \cdot \vec{\mathbf{u}}_T$$

where $\mathcal{A}$ is the elastic tensor, which is his general form is given by

$$\mathcal{A} = \begin{pmatrix} E_{11} & E_{12} & E_{13} & E_{14} & E_{15} & E_{16} \\ \cdot & E_{22} & E_{23} & E_{24} & E_{25} & E_{26} \\ \cdot & \cdot & E_{33} & E_{34} & E_{35} & E_{36} \\ \cdot & \cdot & \cdot & E_{44} & E_{45} & E_{46} \\ \cdot & \cdot & \cdot & \cdot & E_{55} & E_{56} \\ sym & \cdot & \cdot & \cdot & \cdot & E_{66} \end{pmatrix}.$$

With these definitions in mind, we obtain the finite element local matrices

$$
\begin{aligned}
a_{jk} &= \sum_{T \subset supp(\boldsymbol{\psi}_j) \cap supp(\boldsymbol{\psi}_k)} \int_T \mathbf{D}\boldsymbol{\psi}_k^t \cdot \mathbf{D}^t \cdot \mathcal{A} \cdot \mathbf{D} \cdot \mathbf{D}\boldsymbol{\psi}_j \\
R_k &= \sum_{T \subset supp(\boldsymbol{\psi}_k)} \int_T \mathbf{f} \cdot \boldsymbol{\psi}_k + \int_{\partial T \cap \Gamma_N} \mathbf{g} \cdot \boldsymbol{\psi}_k - \int_T \sigma(\mathbf{u}_h) : \varepsilon(\boldsymbol{\psi}_k) \\
&= \sum_{T \subset supp(\boldsymbol{\psi}_k)} \int_T \vec{\boldsymbol{\psi}}_k^t \cdot \vec{\mathbf{f}} + \int_{\partial T \cap \Gamma_N} \vec{\boldsymbol{\psi}}_k^t \cdot \vec{\mathbf{g}} - \int_T \mathbf{D}\boldsymbol{\psi}_k^t \cdot \mathbf{D}^t \cdot \mathcal{A} \cdot \mathbf{D} \cdot \mathbf{DP} \cdot \vec{\mathbf{u}}_{hT} \, .
\end{aligned}
$$

Finally, we can resume the computation of our error estimate using the following flow charts:

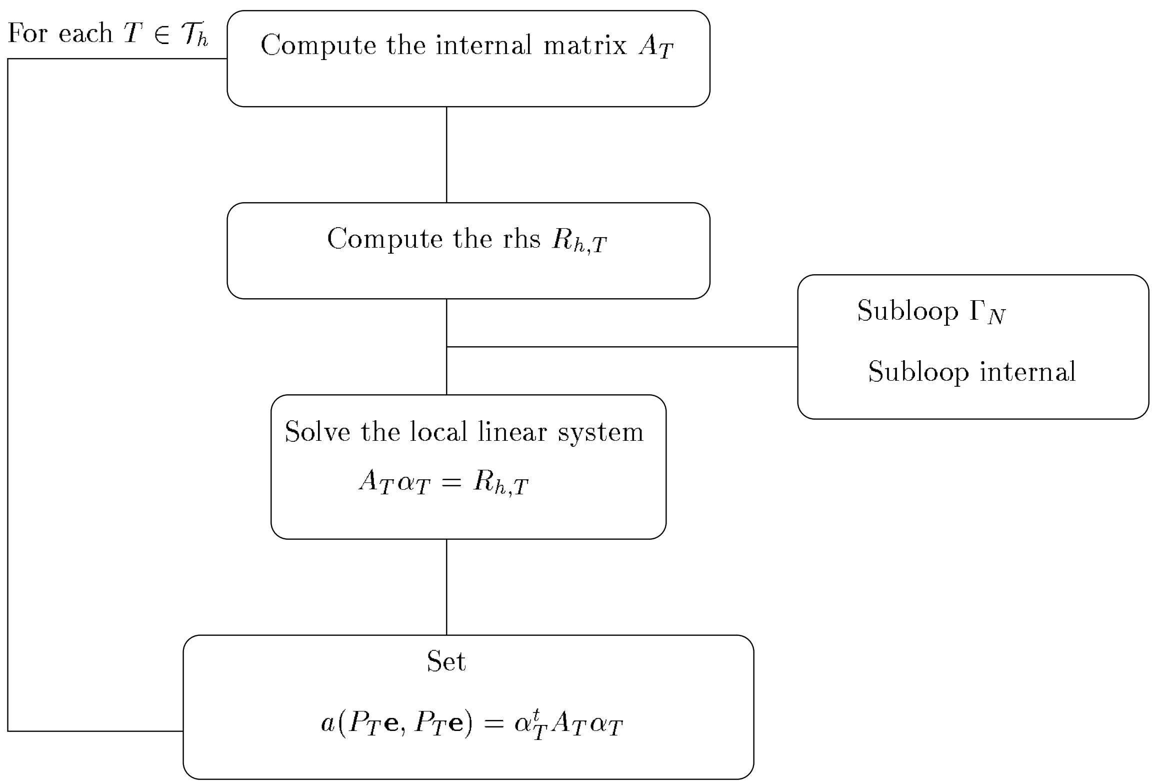

Figure 2: Flow chart to compute the error estimator η for each element

The subloop corresponding to an external face $F \in \Gamma_N \cap T$ is then given by

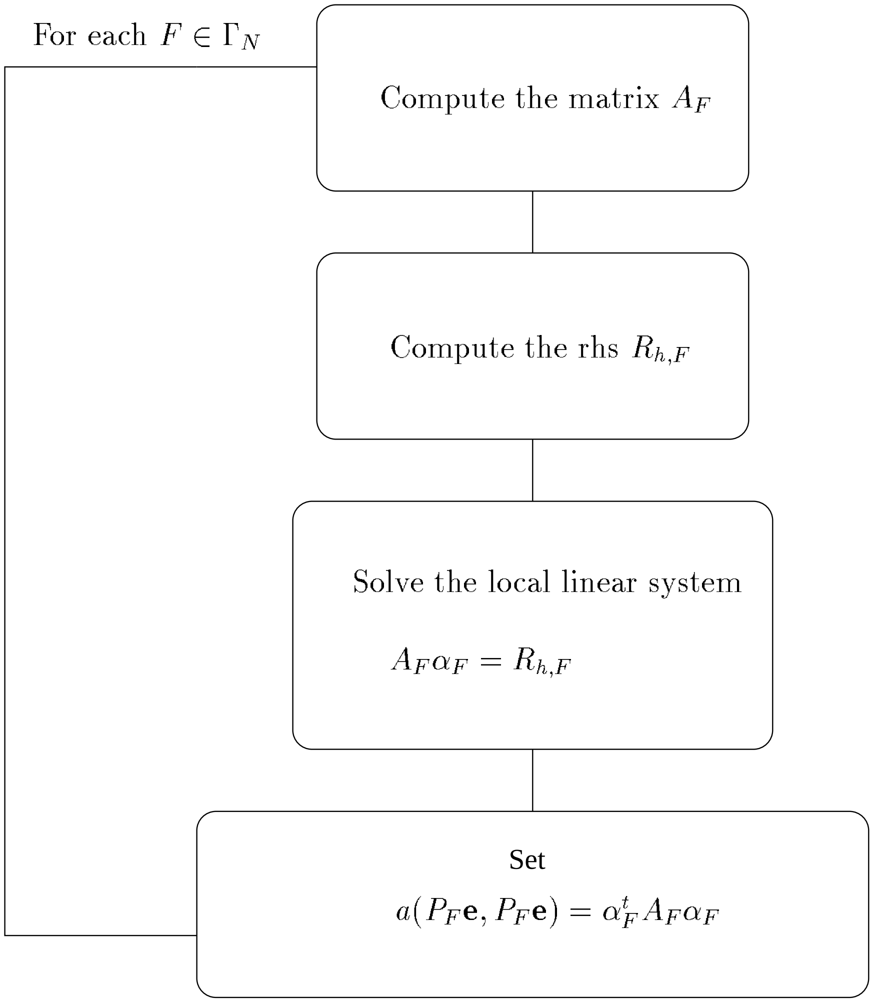

Figure 3: Subloop to compute the error estimator η for each face in Γ_N

For a side F in $\mathcal{E}_h \cup \mathcal{E}_{12}$ the computation is a little more complicated. First, we need to compute the local matrices A_F and the rhs $R_{h,F}$ to form the matrices A_{FF} and the rhs $R_{h,FF}$, this is necessary because each $F \in \mathcal{E}_h \cup \mathcal{E}_{12}$

appears twice when we do an iteration on $T \in \mathcal{T}_h$. Therefore, the subloop corresponding to the construction of the local face matrices is given by

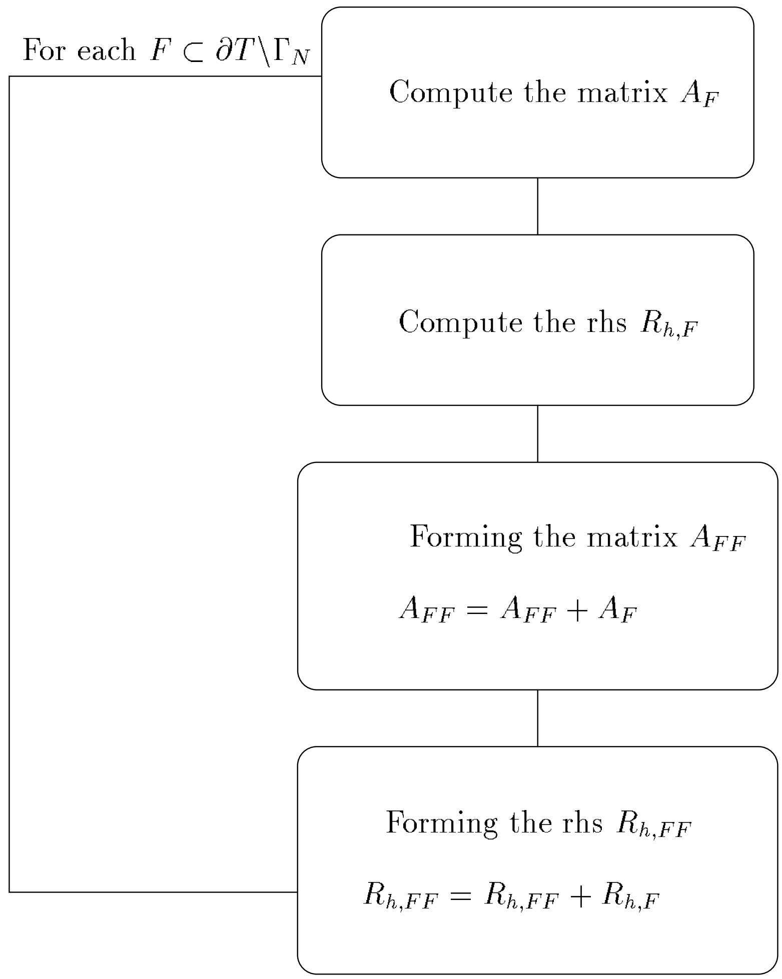

Figure 4: Subloop to construct local face submatrices in error estimate for each face in $\mathcal{E}_h \cup \mathcal{E}_{12}$

The last step for a side (face) F in $\mathcal{E}_h \cup \mathcal{E}_{12}$ is to solve the local system with the matrix A_{FF} and the rhs $R_{h,FF}$ instead of the matrix A_F and the rhs $R_{h,F}$. It is then given by a loop operating on all internal faces

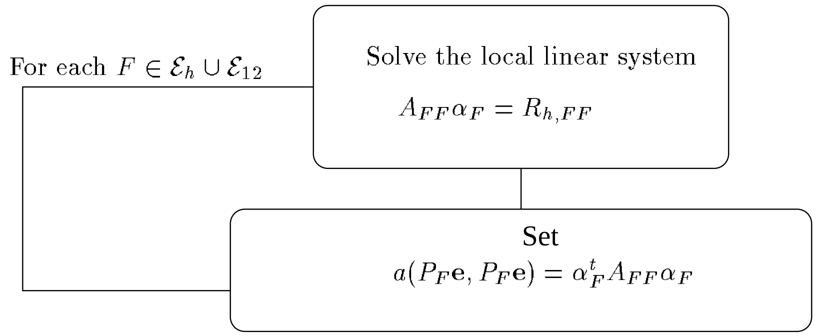

Figure 5: Second step to compute the error estimator η for each face in $\mathcal{E}_h \cup \mathcal{E}_{12}$

6 Numerical examples

In this section we give two examples of the numerical implementation of our hierarchical error estimate. Both examples are calculated using three nodes first order triangles (P1). Here the choice of bubble space is given by the following finite dimensional subspaces

$$\begin{aligned} H_{B_T} &= span\{b_T^* \mathbf{e}_1 , b_T^* \mathbf{e}_2\} \qquad , T \in \mathcal{T}_h \, , \\ H_{B_F} &= span\{b_F^* \mathbf{e}_1 , b_F^* \mathbf{e}_2\} \qquad , F \in \mathcal{E}_h \cup \mathcal{E}_{12} \cup \mathcal{E}_N \, , \end{aligned}$$

where b_T^* and b_F^* were defined in (2.7) and (2.8) respectively.

For each example, we give the initial mesh and a adapted mesh obtained using our hierarchical error estimate. To have an idea of the quality of our adaptation process, we show a cut of the solution obtained by the finite element method in the initial mesh, the adapted one and a very fine reference mesh. The reference mesh is a uniformly refined mesh with 90.000 elements.

Finally, we show a comparison between our hierarchical estimate and two class of residuals estimates. The first residual is the standard one, ie. the error estimate originally derived by Babuska and Rheinboldt [3]. The second residual estimate is a modification of the standard one by the way of introducing a weight in the jump of $\boldsymbol{\sigma} \cdot \mathbf{n}$ on each element interface. This estimate was introduced in the work of Araya and Le Tallec [2]. It's interesting to note that the new hierarchical estimate and the weighted one seem to be numerically equivalent in these two cases.

Remark We follow [9] in order to obtain a optimal mesh refinement procedure, i.e. if ε_0 is the accuracy required by the user, we say that the mesh $\mathcal{T}_h^*$ is *optimal* if its elements number N^* is minimum and it provides a global error ε^* equal to ε_0.

In this framework, for each element $T \in \mathcal{T}_h$, we compute a refinement factor:

$$r_T = \frac{h_T^*}{h_T}$$

where h_T is the size of the element T of $\mathcal{T}_h$, and h_T^* the size of the elements of $\mathcal{T}_h^*$ within the T area (in 2D case).

If no strong gradients appear in the solution, then a priori error estimates indicate that the local contribution to the error should scale like

$$\frac{\eta_T^*}{\eta_T} = \left(\frac{h_T^*}{h_T}\right)^p = r_T^p$$

where p depends on the element type ($p = 1$ for linear elements, $p = 2$ for quadratics elements) and η_T (local error estimate) is given by

$$\eta_T := \{a(P_T\mathbf{e}, P_T\mathbf{e}) + \sum_{F \in \mathcal{E}(T) \cap (\mathcal{E}_h \cup \mathcal{E}_{12} \cup \mathcal{E}_N)} a(P_F\mathbf{e}, P_F\mathbf{e})\}^{1/2}$$

for each $T \in \mathcal{T}_h$.

Thus we have the following minimization problem:

$$\min N^* = \sum_T \left(\frac{1}{r_T}\right)^{dim} \qquad \text{with} \qquad \sum_T r_T^2 \eta_T^2 = \varepsilon_0^2\,.$$

In the 2D case, this problem admits the explicit solution:

$$r_T = \frac{\varepsilon_0^{1/p}}{\eta_T^{1/(p+1)} \left[\sum_T \eta_T^{2/(p+1)}\right]^{1/2p}}\,.$$

The new mesh is then obtained by a metric controlled Delaunay mesh generator [6], constrained to generate local equilateral triangles of size $r_T h_T$.

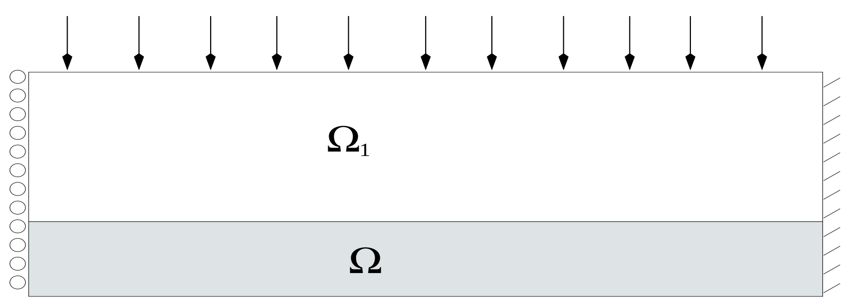

Figure 6: A description of the first example

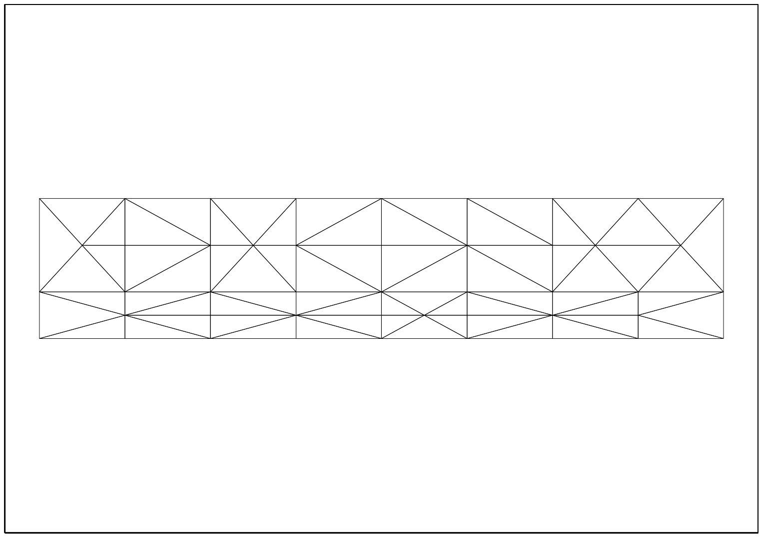

Figure 7: Initial coarse mesh (68 elements)for example 1

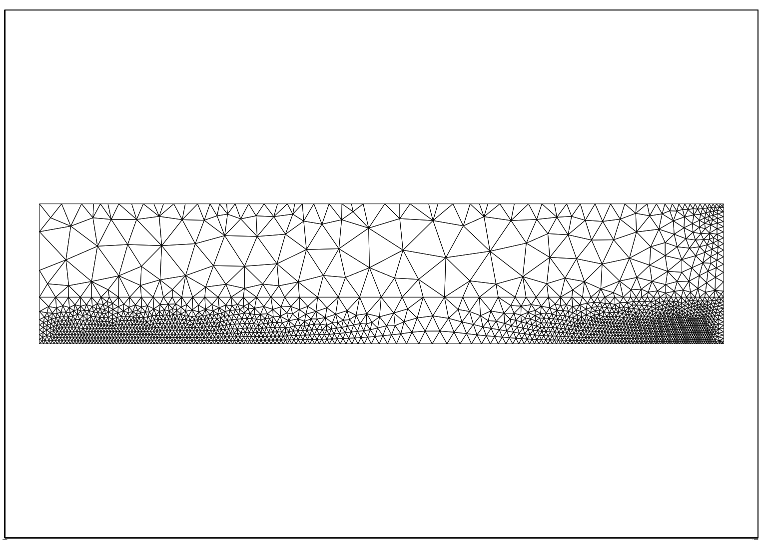

Figure 8: Final adapted mesh (3843 elements)

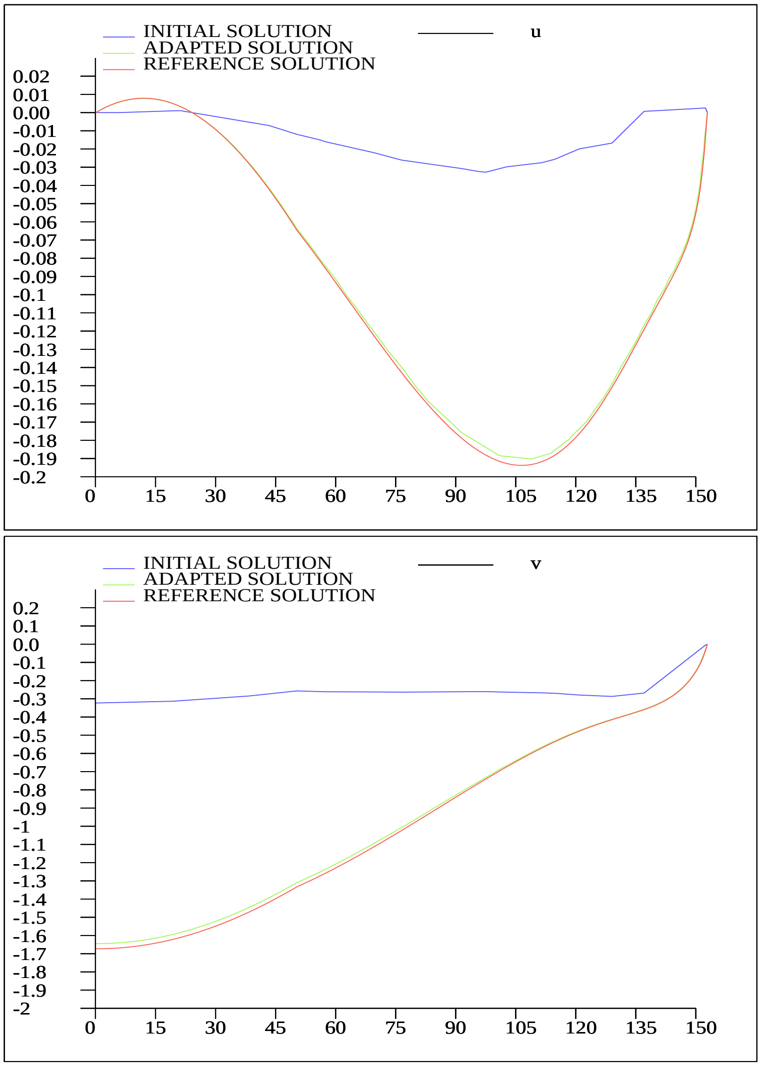

Figure 9: A diagonal cut of the different solutions obtained in the initial, adapted and reference mesh

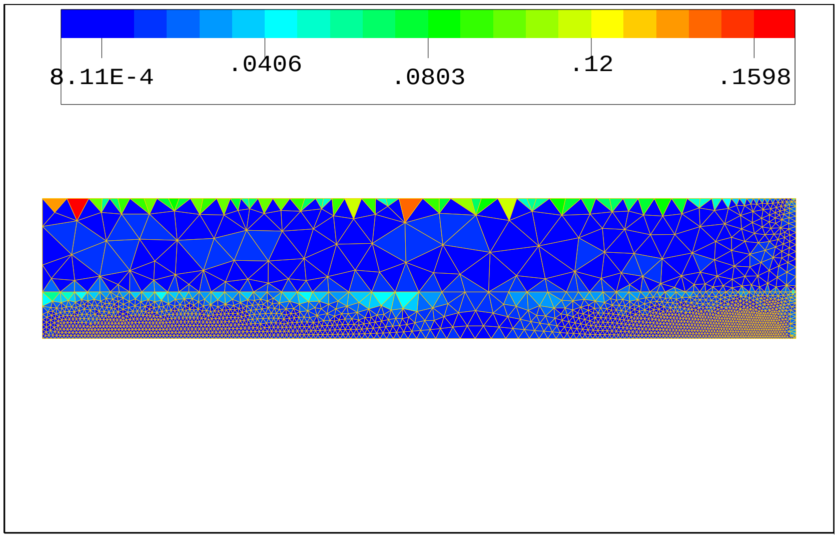

Figure 10: Distribution of η in the final adapted mesh

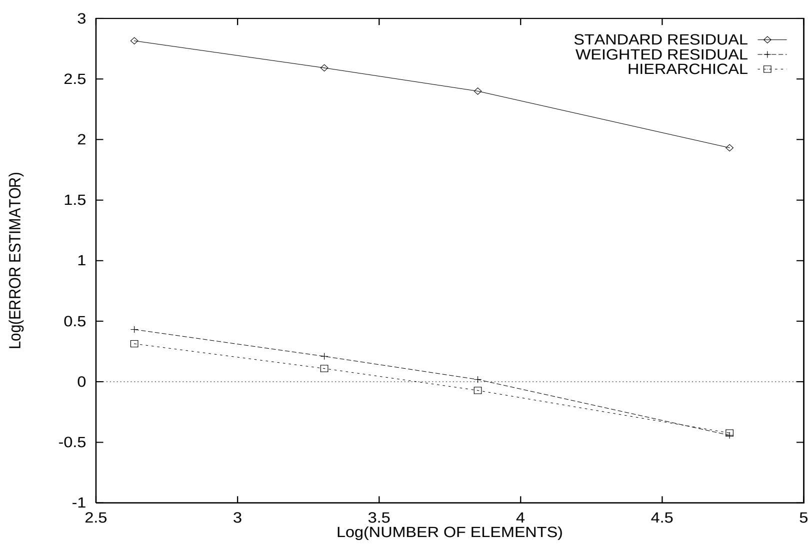

Figure 11: Comparison between different error estimates

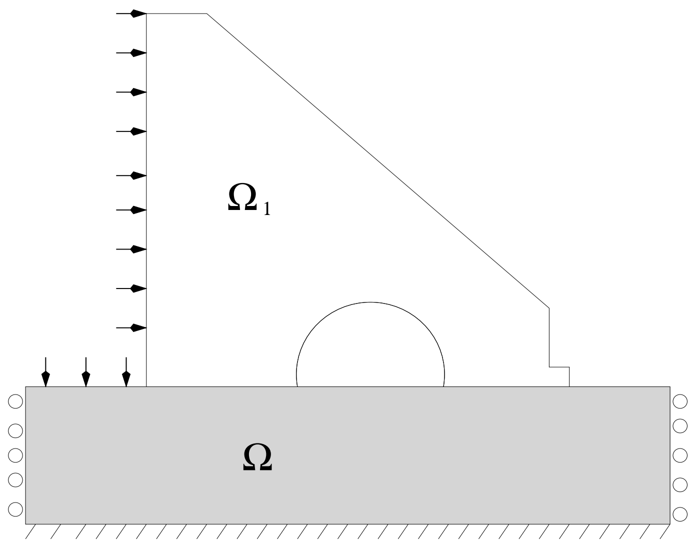

Figure 12: The second example

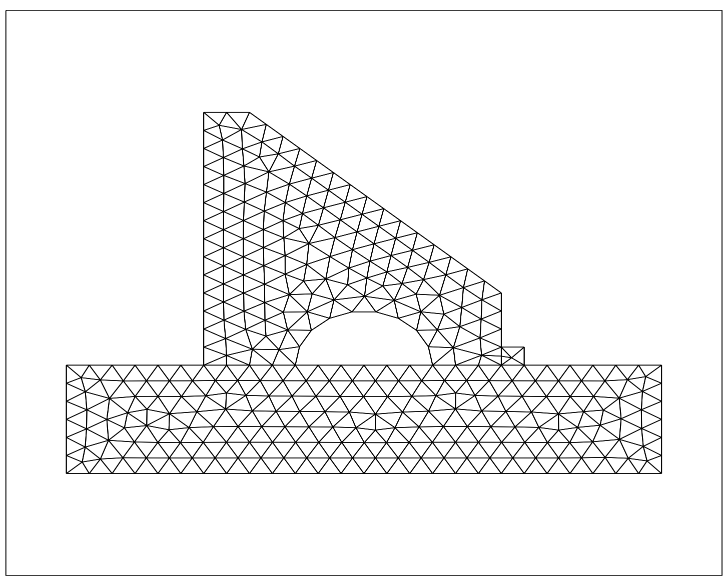

Figure 13: The initial mesh (623 elements) for example 2

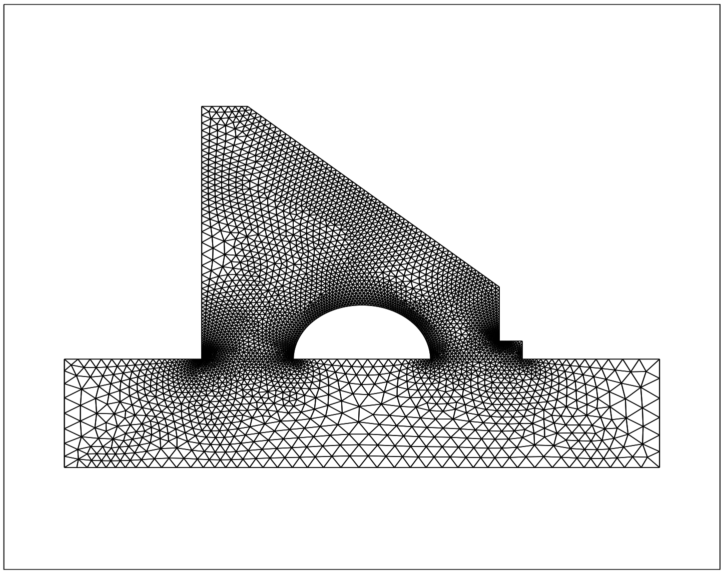

Figure 14: Final adapted mesh (7684 elements)

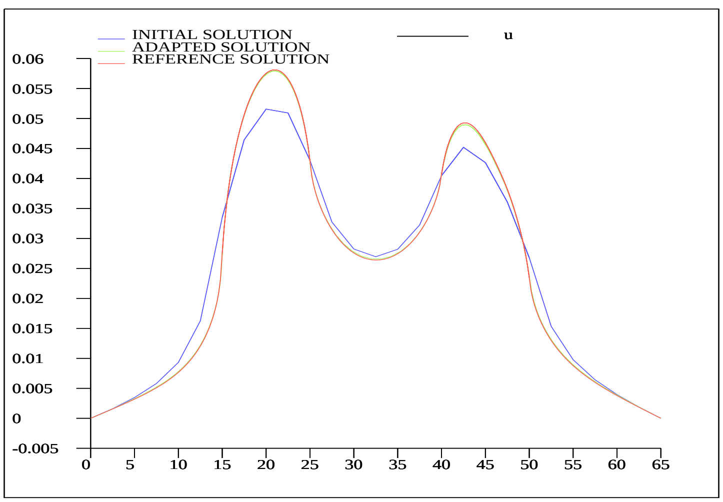

Figure 15: A diagonal cut of the different solutions obtained in the initial, adapted and reference mesh. Adapted and reference solutions are superposed.

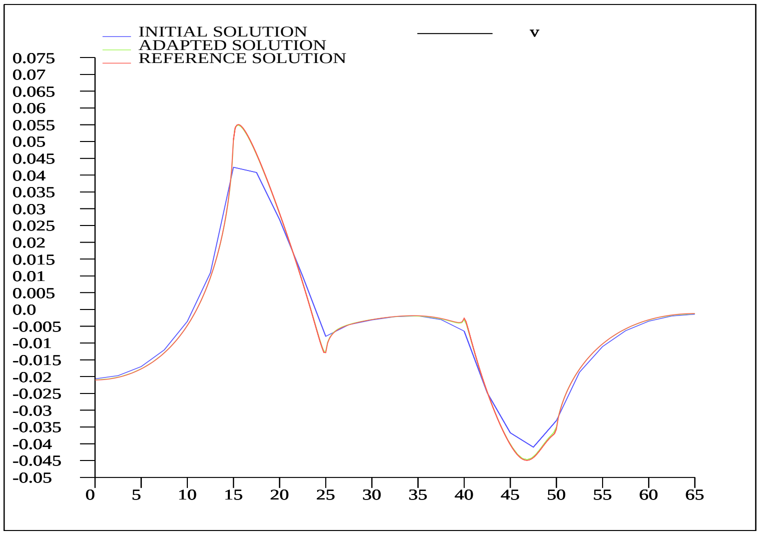

Figure 16: A diagonal cut of the different solutions obtained in the initial, adapted and reference mesh.Adapted and reference solutions are superposed.

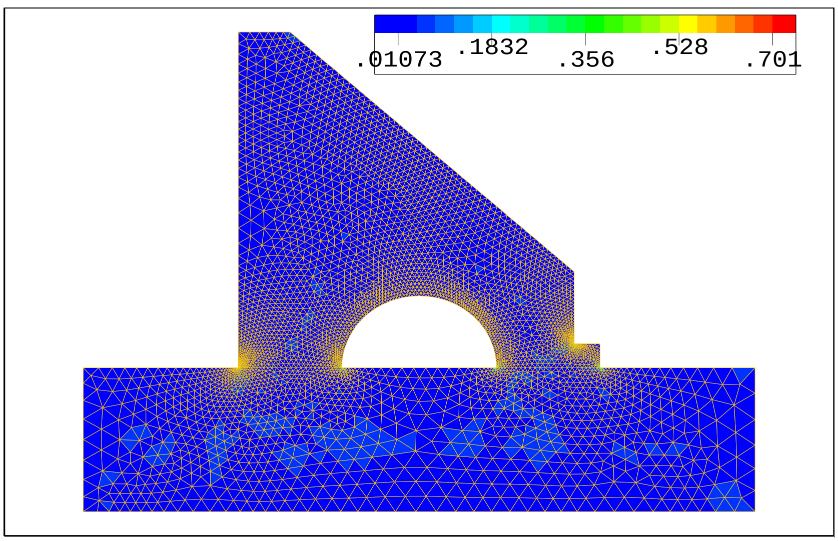

Figure 17: Distribution of η in the final adapted mesh

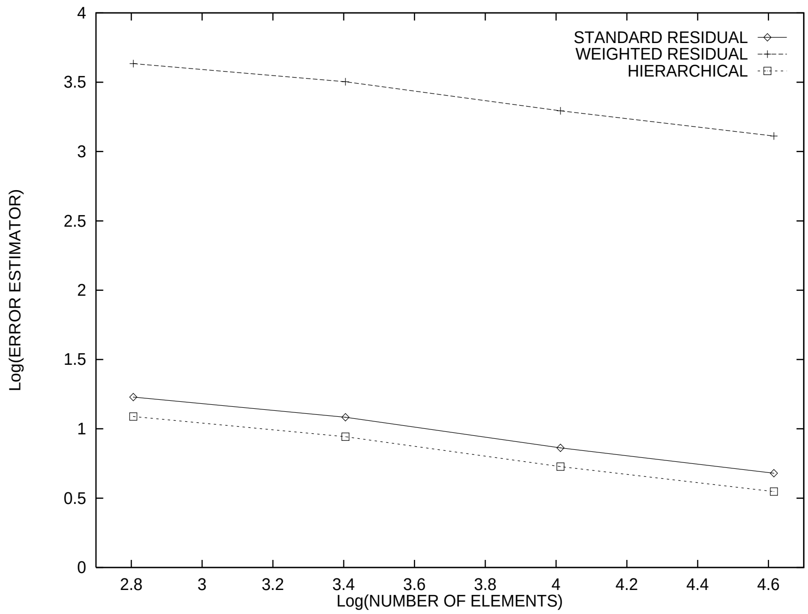

Figure 18: Comparison between different error estimates

7 Conclusion

In this paper, we have introduced a parameter free hierarchical intrinsic error estimate which can be used to improve the quality of a discrete approximation given by the application of the finite element method to an elasticity problem. This estimate is in some way a generalization of a weighted residual estimate introduced in [2].

This estimate is justified by a classical hierarchical error estimate analysis using a saturation assumption and a partition lemma which we have deduced from a inf-sup condition. This LBB condition is more powerful than the two standard assumptions in the sense that it is easier to verify, even for heterogeneous compressible elasticity. In fact, we have proved the optimality of our error estimate independently of coefficient jumps.

For the moment our present work is not adequate to consider the anisotropic case and also this hierarchical approach does not provide a reconstruction of the stress fields unlike the equilibration technique (cf. [9]).

In a forthcoming paper we extend this hierarchical analysis to consider quasi-incompressible materials, the idea here is to obtain an error estimate which will be independent of Poisson ratio ν.

Acknowledgment. The authors wish to acknowledge fruitful discussions with Dr. Barbara Wohlmuth from Augsburg University, and the support of Firtech Calcul Scientifique.

References

[1] M. Ainsworth and J.T. Oden. *A posteriori* error estimation in finite element analysis. *Comput. Methods Appl. Mech. Engrg.*, 142:1–88, 1997.

[2] R. Araya and P. Le Tallec. A robust a posteriori error estimate for elliptic non-homogeneous equations. Technical Report 3279, INRIA, 1997.

[3] I. Babuška and W.C. Rheinboldt. A posteriori error estimates for the finite element method. *Int. J. Numer. Methods Engrg.*, 12:1597–1615, 1978.

[4] R.E. Bank and A. Weiser. Some a posteriori error estimators for elliptical partial differential equations. *Math. Comput.*, 44:283–301, 1985.

[5] F. Bornemann, B. Erdmann, and R. Kornhuber. A posteriori error estimates for elliptic problems in two and three space dimensions. *SIAM J. Numer. Anal.*, 33:1188–1204, 1996.

[6] H. Borouchaki and P. Laug. The bl2d mesh generator: beginner's guide, user's guide and programmer's manual. Technical Report RT–0194, INRIA, 1996.

[7] Ph. G. Ciarlet. *The finite element method.* North-Holland, 1978.

[8] Ph. Clément. Approximation by finite element functions using local regularization. *RAIRO Anal. Numér.*, 2:77–84, 1975.

[9] P. Ladevèze, J.P. Pelle, and Ph. Rougeot. Error estimation and mesh optimization for classical finite elements. *Engrg. Comput.*, 8:69–80, 1991.

[10] R. Nochetto. Removing the saturation assumption in a posteriori error analysis. *Istit. Lombardo Accad. Sci. Lett. Rend. A*, 127(1):67–82, 1993.

[11] R. Verfürth. *A review of a posteriori error estimation and adaptative mesh-refinement techniques.* Wiley and Teubner, Chichester-Stuttgart, 1996.

Unit´e de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unit´e de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unit´e de recherche INRIA Rhône-Alpes, 655, avenue de l'Europe, 38330 MONTBONNOT ST MARTIN
Unit´e de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unit´e de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
http://www.inria.fr
ISSN 0249-6399